Beyond Science and Decisions:
From Problem Formulation to Risk Assessment
Workshop XI

February 18 – 20, 2020
National Institute for Occupational Safety and Health (NIOSH)
Cincinnati, Ohio

Workshop Co-Chairs:
Sabine Lange, TCEQ
Pamela Williams, E Risk Sciences
Neeraja Erraguntla, ACC

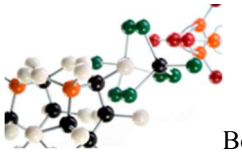
Workshop Coordinators:
Heather King, NIOSH
Christen Williams, TERA
Bethany Hansen, TERA

Rapporteur:
Dr. Bernard Gadagbui,
TERA

Workshop Report
July 24, 2020

Table of Contents

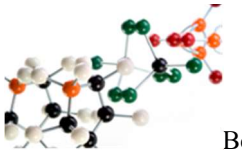
Beyond Science and Decisions: From Problem Formulation to Comprehensive Risk Assessment: Purpose 3
Welcome, Opening Statements and Keynote Address 4
Research Case Studies 12
---Case study 1: Occupational Exposure Banding (EOB) 2.0. Characterizing Risks for Chemicals with Limited Data 13
---Case study 2: Use and Application of Real-Time Exposure Monitoring 20
---Case study 3: Applying Hypothesis-Testing methods to Help Inform Causality Conclusions from Epidemiology Studies 25
---Case study 4: Understanding Weight-of-Evidence of Ototoxicity from Co-Exposures to Noise and Chemicals in the Workplace 42
---Case study 5: Risk/Benefit Methods for Carcinogenicity / Sterilization with Ethylene Oxide as an Example 55
---Case study 6: Risk Assessment Methods of Flavoring in E-Vapor Products 65
Ongoing Activities 93
Appendix 96
---Workshop Information and Sponsors 97
---Background & Purpose 98
---Workshop Objectives 98
---Listing of Research Case Studies 99
---Committees of the Alliance for Risk Assessment 99
---Final Workshop XI Agenda & Purpose 101
---Biographical Sketches of Workshop Co-Chairs, Speakers, Presenters, and Science Panelists 104



Beyond Science and Decisions: From Problem Formulation to Comprehensive Risk Assessment: Purpose

To advance the recommendations in the NAS (2009) report concerning issue identification (problem formulation) and all aspects of risk assessment and management, through selection of illustrative research case studies for further development

Please note that the following report of Workshop XI is only a summary of points raised during the workshop and is not intended to be a complete discussion of all issues.



Welcome, Opening Statements and Keynote Address

Opening Statement by Dr. Michael Dourson, Science Panel

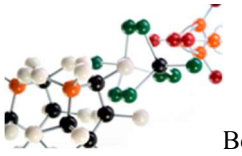
I will give a little bit of history about this whole program, because a lot of us are new to this. The Science & Decisions report was published in 2009 by the National Academy of Sciences. It has a lot of applicability in risk assessment and it talked about new methods and new ways of doing it. Problem formulation was part and parcel of that. Science & Decisions (2009) also recommended using defaults in risk assessment, but also stepping away from these when you have the data. There was a lot of interest in this report and a lot of good ideas, and a lot of ideas that needed to be further developed.

What happened soon thereafter was a petition of Dr. Mike Honeycutt on the Texas Commission on Environmental Quality to the Alliance of Risk Assessment (*ARA*) to form a network of groups to further investigate this Science & Decisions report. The Steering Committee of the *ARA*, a group of folks that oversees this whole operation, agreed to do this. Thus, was formed this Beyond Science & Decisions workshop series. It is simply taking the work of the NAS and adding to it.

After the petition was accepted a project Advisory Committee was formed. The Advisory Committee included folks from different sponsors, four of whom are here today: Sabina Lange from Texas Commission on Environmental Quality, Neeraja Erraguntla from the American Chemistry Council, Dr. Pamela Williams from E Risk Sciences, and myself from Toxicology Excellence for Risk Assessment (TERA). This committee is focuses on progressive risk methods that are fit to a purposeful assessment and that are written in the design of a research case study. This necessitates a well-defined problem formulation and, of course, a risk method often with an example of either a specific chemical or physical agent, chemical mixture, or risk management technique. The Advisory Committee also recommended to the Steering Committee the formation of a Science Panel for this workshop series to give feedback to proffered case studies. This feedback is also enhanced with observer comments.

The overall idea of this workshop series is to be very interactive. The Science Panel provides comments on the research case study, and if the panel judges that the study advances the science, allows it to be loaded on a website that has other accepted case studies. The website gives other researchers an opportunity to check posted case studies to see if one or have comparable problem formulations, allowing them to potentially use the underlying risk methods.

So, with that brief history, I just have several short comments. First, all the presentations and all the case studies we have here are preliminary. So, the authors of the case studies are inviting all sorts of commentary. Second, we have a couple of our panelists who will be coming later, or apparently delayed in traffic. Finally, we are glad you are here. And what I would like to do now is turn this over to Neeraja who is on the risk assessment advisory committee to make a few remarks.



Opening Statement by Dr. Neeraja Erraguntla, Member of the Risk Assessment Advisory Committee

Dr. Neeraja welcomed the participants on behalf of the Advisory Committee. She introduced herself and noted that although she has 20+ years in toxicology and risk management, she is a freshman on this Committee. According to Dr. Neeraja, the main goal is that we have six preliminary assessments. All of them looking into method development or possibly development into case studies later on. The method development is all about how we take the NAS principles and apply them to risk assessment problem formulation to comprehensive risk assessment. She then stated that the Advisory Committee were available to answer questions throughout the time of the workshop.

Keynote Address by Frank J. Hearl, Chief of Staff, National Institute of Occupational Safety and Health (NIOSH) Given by Dr. Christine Whittaker

Introduction by Dr. Paul Schulte (NIOSH). Prior to the Keynote Address, Dr. Paul Schulte, one of the senior leaders of NIOSH, indicated that it was his pleasure to be here and to welcome everyone present to NIOSH. NIOSH is delighted to host this meeting. He pointed out that there was something about the mix here that he thought was going to be surprising to us and about how stimulating these kinds of interactions can be. Dr. Schulte briefly traced all the way back to 3200 BC as the beginning of risk assessment. He then noted that risk assessment has since developed further and that the participants were today going to explore the complexities of modern risk assessment, go beyond what happened when the chemical industry started and the development of occupational exposure limits, which drove a lot of the early thinking about risk assessment. He also noted that the legislation in the 60s and 70s really set the stage for today's work. And that the Science & Decisions (NAS, 2009) report that we heard about earlier is clearly a driver for new developments. He further commented that in the next couple of days we will notice how to advance those recommendations. Dr. Schulte again said that he was excited to be here and that he was also excited for his next task, which was to introduce his colleague, Dr. Christine Whittaker.

Keynote Address: Dr. Whittaker is Chief of the Risk Evaluation Branch in the Division of Science Integration here in the Cincinnati office of NIOSH office. She heads up the NIOSH risk assessment effort and has been the branch chief for going on 15 years. Dr. Whittaker was further introduced as the one to provide an overview or a keynote on behalf of Dr. Frank J. Hearl, who was under the weather and couldn't make it from Washington. After the keynote address, Dr. Whittaker was also going to present the NIOSH case study.

Dr. Chris Whittaker briefly described why NIOSH assesses chemical hazards in the workplace, including the steps that are involved in its risk assessment process, and where NIOSH is going

with risk assessment in light of recent events such as the new Lautenberg Amendment to the Toxic Substances Control Act (TSCA). According to Dr. Whittaker, about one third of U.S. workers are exposed to chemicals. This translates into more than 50 million workers being exposed to hazardous chemicals daily. In 2016, the Bureau of Labor Statistics reported there were 12,480 non-fatal lost time injuries due to chemicals. This grossly underestimates the toll that chemicals take in estimates ranging from 400 - 500,000 diseases that are caused by chemical exposures every year. About 2-8% of cancers are believed to be caused by occupational exposures. According to Dr. Whittaker, NIOSH derived its authority from the Occupational Safety and Health Act of 1970, in which NIOSH was directed to develop exposure levels that are safe for various periods of employment, including, the levels at which no employees will suffer impairment of health or functional capacity. So, in the early days of NIOSH risk assessment in the late 1980s, NIOSH took on radon, benzene, and glycol ethers and then moved into coal mine dust and metalworking fluids and noise. The progress was slow in the beginning, but things picked up a little bit in the 2000's, with NIOSH publishing significantly more quantitative risk assessments. The focus became more on looking at methods, and how such methods can be used to make quantitative assessments of the risks for different kinds of chemicals.

NIOSH began providing Recommended Exposure Limits (RELs), which are generally time-weighted eight-hour average limits and also Risk Management Limit for Carcinogens (RML-CA), which were proposed in NIOSH's cancer policy published a couple of years ago. NIOSH is developing some of the first RML-CAs. NIOSH's primary publications are in Criteria Documents and Current Intelligent Bulletins (CIBs). The Agency's guidance on chemical exposures and chemical exposure limits are in the NIOSH Pocket Guide to Chemical Hazards. Regarding NIOSH's risk assessment process, their focus is on determining the relationship between the predicted occupational exposure and the adverse health effect, whatever that health effect might be. And, in the very near future, hopefully, this spring, NIOSH practices in occupational risk assessment document will be published. Dr. Whittaker showed a mockup of the cover and hoped that's what it is going to look like. This document describes the current practices in the NIOSH occupational risk assessment and provides the underlying philosophy and support for NIOSH RELs, RMLA-CAs, Short-term Exposure Limits (STELs), etc.

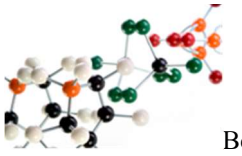
Dr. Whittaker also briefly described the NIOSH occupational risk assessment paradigm, starting with hazard identification, which determines the kinds of health effects caused by a chemical. This is followed by a response analysis and risk characterization, all of which is informed by the mode of action of the chemical. This paradigm then feeds right into a risk management and risk management recommendations. NIOSH uses engineering controls to control hazards or personal protective equipment, and other administrative controls. NIOSH does not have a separate step for exposure assessment the way EPA does exposure assessment in its risk assessments. Instead, NIOSH compares the exposure to a target risk level. Exposure assessment is an integral part of NIOSH's occupational epidemiology studies, and exposure assessment is considered in developing the engineering controls. In effect, NIOSH does not do a separate comparison of what workers are exposed to now, or what the risk levels might be. Everything is put together and feeds into both risk policy and risk science.

The current NIOSH risk assessment priorities are propane, manganese, lead and others. NIOSH

is also working on chemicals that are at earlier stages. Now, enters the Lautenberg Amendment of TSCA passed in 2016 by bi-partisan Congressional agreement. This amendment changed EPA's mandate for how they are doing risk assessment. The phrase that affects occupational risk assessment is in Section 2605 under the requirements, where the requirement is to integrate and assess available information on hazards and exposures for the conditions of use of the chemical substance, including information that is relevant to specific risks of injury or environment and information on potentially exposed versus susceptible populations. This includes workers. So, a lot of the risk assessment that have been coming out of TSCA have focused on worker exposures and there is a very good reason for this. Worker exposures are generally higher. It is easier to make the case on whether there is unreasonable risk or not. If you have a worker exposed to higher levels, this phrase has really changed the landscape as far as occupational risk assessment goes. Other changes that are important to NIOSH are that the EPA was mandated to make a determination of unreasonable risk or no unreasonable risk for its chemical assessment. It evaluates conditions of use and all of the scenarios of how a chemical is used. It has rigid statutory deadlines to complete risk assessment. It has designated both high and low priority chemicals. EPA has identified 20 high-priority and 20 low priority chemicals that are coming up in the next batch. That is a lot of chemical risk assessment going on, and a lot of focus on workers. There are a lot of familiar chemicals in EPA's first 10 chemicals, including asbestos and methylene chloride. EPA is at various stages of completion these risk assessments, going through peer review and public comment. Methylene chloride and one other are completed. And some of them are in earlier phases. Simultaneously EPA is also proposing their next 20 high-priority chemicals. These will have the same three-year or three- and half-year deadline as the first 10 did. In the second batch, EPA is looking at a lot of solvents and some other chemicals, including formaldehyde and others as well as old favorites. The high-priority chemicals pose some challenges for EPA. Although the next 20 are not finalized yet, they are the low priority chemicals since they do not seem to be particularly hazardous from an occupational point of view.

In light of this, new TSCA mandate, what does NIOSH risk assessment look like in the age of Lautenberg Amendment? NIOSH is likely to continue to conduct fewer single chemical risk assessments and focus on some other pieces of the risk assessment pie. There are lots of occupational risk assessments to go around. NIOSH will increase its focus on acute catastrophic hazards and, maybe, assess chemicals with limited data. A lot of the chemicals that EPA is looking at right now have significant amounts of data although some may have limited data to begin with. NIOSH is likely to be looking at how it assesses all those thousands and thousands of chemicals without data. NIOSH would increase its focus on endpoints such as irritation, which although it does not appear to be a severe endpoint, it actually is the endpoint that is the basis of half of the RELs in the NIOSH Pocket Guide.

It will be useful to integrate some of EPA’s TSCA risk assessments with the NIOSH guidance and take their risk assessments and move them forward by working together as one government. NIOSH can expand its Occupational Exposure Banding (OEB). That is what the NIOSH case study is going to be talking about and moving that to the next level: real-time monitoring and risk assessment. We will also have a presentation on the issue of what happens in very short, very high exposures. We typically think of risk assessment in terms of 45 years of exposure or



eight hours of exposure. It is equally important to know what happens if your exposures are five minutes or six seconds as these short-term exposures can present their own problems.

Beyond chemical risks, NIOSH is also conducting biological risk assessment. Another effort is looking at psychosocial risk assessment and applying some of the things we have learned in chemical risk assessment to other types of hazards in the workplace. For irritation and occupational risk assessment, 50% of the RELs, translating to well over 300 chemicals in the Pocket Guide are based on irritation effects. The second most common health effect is cancer. This means that irritation effects and cancer are the two things that are important. Irritation is not just a health issue, but it is also an economic issue for employers, because if a worker can't see, because their eyes are watery and they can't work effectively because they can't breathe, then they are not going to be working as much. Since cancer takes a long time to develop, this is not necessarily going to affect the activity like irritation issues may. Currently, there is no standardized method for assessing irritation endpoints and coming up with RELs. There have been some proposals in the literature and NIOSH is looking into those issues and researching the use of those methods.

Immediately Dangerous to Life and Health (IDLH) values

NIOSH has also IDLH values. Irritation is very important in setting IDLH values, because if you have an irritating chemical that makes it so you can't see or can't breathe and can't escape the environment, the toxic environment is considered in setting IDLH values. The method of setting IDLH values has varied over the years since we have been setting IDHL's.

Short-term Exposure Limits (STELs)

Irritation effects are the major endpoints considered in setting STELs. NIOSH scientists are also building off earlier research in mode of action (MOA) for irritants and looking more into that literature and seeing what they can say about classes of chemicals and their different MOAs. Some animal studies are going on that look at improving the RD50 method to make it more humane for the animals and also more reproducible. Histopathology data are being compared with RD50 data to make sure people understand what is meant by saying a chemical is an irritant. Time assumptions are being evaluated as well. For that reason, 15 minutes became magical for STELs when that was the average time for the analytical methods, say in the 1970s. With real-time instrumentation, 15 minutes is no longer a magical time frame. Therefore, NIOSH is looking more closely at issues of exposure duration.

Occupational Exposure Banding (OEB)

NIOSH recently published the NIOSH OEB process for chemicals and risk management. This has an associated e-Tool online and the automation function of the e-Tool is being improved. Currently, Tier I is fully automated. A little more about the OEB will be discussed in the NIOSH case study. NIOSH is also looking into expanding the understanding of how to make thermal exposure banding and make it more useful for emergency responders, who are most concerned with acute exposures. The case study will also involve brainstorming about Banding 2.0 and

what that means. Another consideration is the potential to study chemicals with very little data available.

Toxicology Testing in the 21st Century (Tox21) and Occupational Risk Assessment:

NIOSH statisticians are looking at predicting dose-response curves based on machine learning curves. There are a lot of efforts going on around the world with researchers trying to predict point estimates of toxicity. Dr. Wheeler of NIOSH has looked at not only predicting a point estimate but approaches to predict the entire dose-response curve for the types of Tox21 data such as high-throughput screening (HTS), and initial evaluations show that it will work. Dr. Wheeler's investigations shows it is possible to predict the shape of the dose-response curve and it can be done very reliably with Tox21 HTS data. More will be learned going to more sophisticated or more advanced types of toxicity testing. This is one area where a lot more research is needed, and this is currently being done. As one gets higher levels of toxicity testing into whole animal and full lifetime bioassays, the dataset gets smaller and smaller. Instead of 10,000 or 100,000 chemicals or assays, one is down to hundreds. Thus, validating the reliability of the method becomes a problematic. That is a problem that NIOSH has not overcome yet. As time is of the essence, recommendations are tied to time. Eight-hour time weighted averages, 15-minute exposures and current consideration is that time measurement is not to exceed a 30-minute exposure. That doesn't mean it would be safe for 30 minutes but that one might survive after 30 minutes, but the point is to get out as soon as one can. However, what does it mean if you exceed an IDHL value for five minutes, or one minute, or six seconds? Is that a hazard, because with all the real-time monitoring that is what we are seeing now? Is that a problem? That is one of the issues NIOSH is confronting.

Biological Risk Assessment:

There are some projects looking at transmission of infectious diseases in the workplace, including people to people, animals to people, modeling surface contamination, and air transmission in confined spaces like aircraft. There are many similarities to chemical risk assessment, but there are some challenges as well as some differences that we need to take into account.

Cumulative Risk Assessment

Cumulative risk assessment is part of the total worker health initiative at NIOSH. This comes out of consideration of things like mixed exposures at work and expanding that to looking at personal and occupational risk factors. This includes risk factors you would confront in your environment, in society, as an individual, along with your occupational risk that you face every day. And, right now NIOSH is at the point of developing frameworks in how to study this and look at the different pieces and their importance. The future of this work depends on the changing nature of work, the changing nature of the workplace, increasing presence of robotics in the workplace, emerging hazards and the changing nature of the workforce. When you consider older workers, a bigger economy, etc., it is equally important to know how all these would impact risk, training, exposures and things like that. The future of work would involve a

30-hour work week, but that might not be official. So, in the age of the TSCA Lautenberg Amendment, cumulative risk assessment is changing. We may well have less emphasis on individual risk assessment for individual chemicals, but we will continue to have some going on in those areas. More emphasis will be placed on the other impacts of chemical exposures and more emphasis on the complex challenges in risk assessment with limited data, a changing timescale that we are looking at, and complex exposure patterns that influence the risks we face every day, at work and in our lives.

Dr. Whittaker thanked the audience for their attention and welcomed them to the National Institute for Occupational Safety and Health.

Q&A:

Q: Is there any effort to look at new approach methods for hazard characterization, rather than animals? EPA committed to no animal testing by 2035

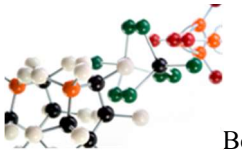
Dr. Whittaker: Yes, part of the Tox21 effort is looking at chemicals with very little data available and alternatives to animal testing. However, occupational epidemiology and animal testing are the gold standard right now, so it is hard to get past those for a deep understanding of chemicals right now.

Q: It looks like Dr. Whittaker is underplaying the strengths of NIOSH, in particular. EPA absolutely needs NIOSH input, particularly with respect to their expertise in analytical methods and some of the understanding NIOSH has in terms of exposure assessment to better characterize EPA's use cases for Lautenberg. There is a huge opportunity for partnership in the age of Lautenberg for NIOSH and EPA to better work together, because EPA can't characterize the use cases without NIOSH

Dr. Whittaker: NIOSH currently has a large effort in supporting EPA's risk assessments under TSCA with both review and providing data and expertise in occupational risk assessment as well as in occupational epidemiology.

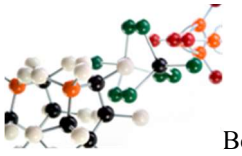
Q: With regard to NIOSH developing new OEL type factors for cancer, could you give us a little preview of what the methodology you are thinking of for the RMLCAs and how you consider that in carcinogens?

Dr. Whittaker: The methodology is the same risk assessment and you end up with a RMLCAs, as the previous REL. But the interpretation is different. Before 1995, anything that was considered a carcinogen, NIOSH recommended as low as feasible, but without assigning any numeric values with that. In 1995 and since, NIOSH switched to a more quantitative risk assessment approach. The RMLCAs is an acknowledgment that an REL for a non-carcinogen is not the same thing as an OEL for a carcinogen. It is a starting place. The goal is to get exposures as low as feasible. It is a kind of a philosophical change more than a methodological change.



Q: The question is about home-based occupational exposures and off-site exposures in the age of the TSCA Lautenberg Amendment. The definition of going to work is changing rapidly in this country, not just in terms of offices, but small businesses or extensions. Many things are being done now in remote locations. Also, there is the concern of off-site exposures, such as painters and people who are not in a building. And, sort of an offshoot from the whole thing which is, take-home exposures which, could be interpreted to mean the secondary exposure from the occupational exposure. In this situation, one is exposed at the work place and one wears those clothes home, walks home with those shoes and it turns out that the highest exposures are in the household, not back at the place where people were being "protected". So, is NIOSH considering any of these scenarios in its study?

Dr. Whittaker: NIOSH has had a long history of looking at take-home toxins, which goes back to the asbestos work and the lead work. In fact, NIOSH has a group now that is specifically looking at some issues of lead take-home toxin from lead workplaces. There is also a concern about people who don't work in traditional workplaces, like people who go out and renovate homes. NIOSH has always had efforts going on to characterize their risks as well. As far as how workplaces are changing and that is moving more towards a dispersed workforce, that is something being considered in our future of work initiative quite strongly. How the workforce is changing, and what does the workplace look like are some of the issues being looked as well as how to characterize work. So, NIOSH has a lot of different efforts, but Risk Evaluation Branch is concerned with how to characterize risks quantitatively to some of those issues.



This page left intentionally blank.

Research Case Studies

Case study 1: Occupational Exposure Banding (EOB) 2.0. Characterizing Risks for Chemicals with Limited Data

This case study was presented by Dr. Christine Whittaker of NIOSH. The presentation is available on the *ARA* website. The purpose of this case study is to explore how NIOSH can create a consistent and documented process to characterize chemical hazards, so that employers could make well informed decisions for those chemicals lacking an occupational exposure limit (OEL).

Discussion and Comments from Panelists and Audience

(1) **Question/comment**: How do data gaps get addressed in the current banding? What is coverage of data points? What is minimum data base? Do you incentivize collecting data, if for example, your minimum when the real critical toxicity of a chemical is missing?

Response by Dr. Whittaker*:* Currently, data gaps are addressed in the EDS/determinant score. The way we looked at that was if a chemical had 75 LD50 values and no 2-year bioassay, we would say that is insufficient data. We wanted to focus on data that characterized more towards full lifetime exposures, so we have a better idea of what the hazards are over the lifetime. Two-year bioassays, repeat exposures, and reproductive toxicity would be weighted more than acute studies. Basically, you need at least one long-term assay (cancer, repro, or target organ) in the database for each chemical to meet our sufficiency data requirement. Acute toxicity, skin irritation, and eye irritation only contribute 5, whereas longer-term studies contribute 20 or 30 depending on what the data are. There is a minimum of 30 that you need. It does not matter how many studies you have in one endpoint. If you have 75 studies in one endpoint, for example, LD50 studies, you still only get a 5. You do not get 75 times 5. That is all it says. It is the presence of data that has been vetted by some agencies. Regarding missing data, we are trying to make sense of what data are available. This does not really incentivize collecting data.

(2) **Question/comment**: Do you incentivize collecting data? Your scoring criteria may indicate you meet the minimum score of 30 but you could be in a situation where the real critical toxicity for the chemical is missing and you are still banding it.

Response by Dr. Whittaker*:* We are trying to make sense of what data are available. Our process does not really incentivize collecting data. If a chemical does not meet the sufficiency, you cannot band if so then that would incentivize somewhat. There is a concern if information on the real critical toxicity is missing, but we are always in the position of doing the best we can with the information we have.

(3) **Question/comment**: Comment. I really applaud you for making the banding process automated. However, I worry about curation of chemical structures. In EPA, we found a lot of SMILES that were incorrect for chemical structures. I would advocate for using a curated database like EPA's dashboard for chemical structures because the physicochemical properties of the chemicals were

missing in what you described. I am assuming that your categories for particles assume that they are inhalable or respirable, but that is not articulated. You could have a huge amount of particle exposure that will not get inhaled, so it does not need to be in that particular exposure band. That is an opportunity for communication. Likewise, with respect to gases, if the chemical structure is known, one might have more concern for respiratory or irritation effects. I have a very different reaction to coverage for certain endpoints for reactive gases. There is also the need to include chemical structure and likely mode of action of gases categories.

(4) **Question/comment**: Comment. I really appreciate the tiered approach. How are the qualitative cut points developed for each of the endpoints in Tier 2 bands?

***Response by Dr. Whittaker**:* A lot of time was spent on developing the cut points and it depends on the endpoint. We looked at current distribution of chemicals with OELs to see where they fall for different endpoints. The quantitative cut points were based on distribution of chemicals with OELs. It is not a perfect system and certainly the OELs that exist are not perfect either. A lot of judgement was used, and we tried to base them on where we would expect them to fall within a distribution. Ground-truthing or validation of the process was done using diverse industrial chemicals prior to publishing the method. The process is not perfectly reproducible, but reasonably consistent. A journal article is going to explain that a little bit more in depth.

(5) **Question/comment:** Comment. Going back to data gaps. It may be helpful to consider physicochemical properties and descriptors, as well as quantitative structure activity relationship (QSAR), which is most useful for a limited range of endpoints (e.g., cancer, genotoxicity). Considering all of these descriptors will add greater confidence in the banding process. A weighted set of characteristics have been developed in the complex hazard tool for the Domestic Substances List (DSL) process for Canada (very generally). These types of tools need to be designed for purpose, and while the availability/complexity of QSAR has increased more recently, its application is still largely for genotoxicity and cancer. For read-across, more data on like-chemicals are needed. You could consider hierarchical weighting and different lines of evidence, combining data, QSAR analysis and read across. It seems important to be explicit about how you are weighting different information sources and why; it normally depends on what is important in your particular system based on your previous experience with more extensive data sets. You mentioned that it is a challenge to develop a cohesive system for use by a broader audience. Making the process user-friendly is key. For example, you need clear decision rules for QSAR – but still, application is generally difficult without expert judgement. European Food Safety's (EFSA's) library of models under development can be considered as a source.

(6) **Question/comment:** Comment. There is not enough data for a lot of chemicals. High throughput data being collected under ToxCast can be used for this purpose. There is opportunity to bring ToxCast and QSAR/read-across together. A range of HTS tests capture different endpoints. Tests can identify biological targets based on heat map. ToxCast can thus be used to identify most sensitive targets. The most sensitive heat on the map is assumed to be equivalent to the dose in vivo. EPA is developing generic physiologically based pharmacokinetic (PBPK) models to translate that *in vitro* concentration to in vivo exposure. This in vivo dose can then be used

possibly in Tier 2. You have to have expertise in terms of understanding those interfaces. That will take a lot of work but it's a real opportunity.

(7) **Question/comment:** Comment. Picking up from the previous panelists, there are certain ways to categorize chemicals based on physico chemical properties - particles, gases. It is important to consider *in vitro* to in vivo extrapolation. The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) has shown that *in vitro* skin sensitization method is a better predictor than the Localized Lymph Node Assay (LLNA) and should be used. NIOSH should pay more attention to portal of entry effects. ToxCast is not good yet for inhalation and respiratory effects. The key issue here is ground truthing and case study ground truthing with worker data would be helpful. The EPA in Research Triangle Park (RTP) is looking into expanding ToxCast domain to include aerosols, Category 1 and 2 gases. However, the challenge is characterizing flow rates and other related dosimetry aspects for *in vitro* exposure systems. Currently, a document is in preparation on reporting standard to reduce variability and the need to measure what is getting into the humans and into the cells in order to scale up to *in vivo* conditions.

(8) **Question/comment:** Since most of the chemicals have little exposure, it is important to know which chemicals need attention for workers. There is the need to prioritize chemicals in question. There are perhaps 85,000 chemicals in production but the vast majority of those, probably 50,000 of them are likely to have exposures that are essentially going to be a very small volume in use. In the case of industry, for example, when you run into those situations and if you are in production of this kind of chemicals and only making small quantities, you basically tell the business who is responsible for the chemical that they have two options. You can assume a worst case and that means that 2 or 3 of their workers may once in a year will have to have full personal protective equipment (PPE) or they can invest in this particular toxicology test , then they could move to this relief. You need incentives for data collection.

(9) **Question/comment:** How do you determine your priorities since this is not clear from the case study? It doesn't appear that there is much potential to prioritize on the basis of exposure because the system is completely hazard based.

***Response by Dr. Whittaker**:* We do not prioritize the chemicals. The method is hazard-based and so not prioritized based on exposure. We have just developed a tool that can be applied to all chemicals. Exposure is not considered but just the hazard.

(10) **Question/comment:** There is an opportunity to enhance the banding process if exposure assessment principles are used to complement existing framework. The employers had a "treasure trove" of information. They are not randomly using chemicals. They know why the chemicals are used and understand how much is used (e.g., as surfactants, paints, colorants, to absorb UV radiation, etc.) They know they will use a lot or little and what the processes are to get you from the ingredients to the end product or whatever is moving outside. Because of Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), databases are being developed in Europe to get to the heart of why a chemical is used in a product. Although probably not appealing, you could completely walk away from any toxicology and make a

complementary system where you have contaminants that speak to profiles of components that speak to profile of exposure. This is not a standalone, and it is not suggesting this replaces anything, but it certainly could complement your existing profiles by modifying or weighting certain kinds of toxicology data so that it moves up to a higher category. You could create a use and condition profile for a lot of chemicals by approaching the people in the industry and asking why they use these chemicals or selling chemicals for specific ingredient use. To the extent you can, you can categorize the chemicals to create some kind of system that would complement the existing framework. Sometimes the treasure trove is cheap data. Set up the use and condition profile the way you did with the toxicology and see if you can combine it or use that for guidance.

***Response by Dr. Whittaker**:* TSCA people would really love that because they are always looking to define the conditions of use because that is under their mandate. The other thing is that OEB came out of the control banding paradigm and OEB was kind of seen as the 1st step. This is looking at the hazard part. Marrying it to exposure considerations and how people use the chemical was always envisioned as kind of the next thing. As far as to make it practical for employers, the employer needs to interpret the number from the band in light of what they are doing in in their company. The kind of back engineering and coming up with a database, would be very cool.

(11) **Question/comment:** I was wondering how banding relates to control. We've been profiling data on uses of chemicals for some time and while it can be labor-intensive, this information is increasingly available publicly and from a large number of sources and now it's being compiled, so it seems like an extremely valuable resource to additionally consider.

(12) **Question/comment:** Comment. Picking on the comment made earlier regarding incorporating exposure part of the assessment, consider building more exposure as a part of your overall opportunities that are available for the situations where chemicals have very little information on them. The concept of the threshold of toxicological concern (TTC) comes to mind. This concept only works if you understand exposure.

(13) **Question/comment:** Comment. At a meeting last month, the US Food and Drug Agency (FDA) was putting forward a modified version of the threshold of toxicological concern (TTC) approach. They have already dug into it and one could find several publications online. You can take the concept when you absolutely have nothing much on the chemical more than structure. This does not supersede what you have, but when you have nothing but structure, may be that is something to look at.

(14) **Question/comment:** Comment. I just want to reinforce emphasis on physicochemical properties because not only is it the exposure but then you have to translate that to internal dose, and that depends on physicochemical properties, which is why some of these OEB would be less concern to me than for volatiles for example. The physicochemical properties can help inform dosimetry and coverage considerations. It could also be reflected as well in the TTC physicochemical. Your lowest hanging fruit is to superimpose consideration of the likely categories of physicochemical properties. As mentioned already, they are using them for certain reasons.

They know what those properties are. One clarifying question though is that once a band is established, do you go back and update them? Do you encourage people to stay up to date?

***Response by Dr. Whittaker**:* We do no band anything. It is a tool. We encourage people to revisit it, but we obviously do not have control whether they do or not.

(15) **Question/comment:** Thinking about the user, you mentioned an employer. I am just wondering if you can give us a mental model of the employer as to whether they are a small employer or medium-sized or if they have industrial hygiene staff.

***Response by Dr. Whittaker**:* Originally, we were trying to come up with something that would be useful for small employers who would not necessarily have industrial hygiene staff, but this tool is not quite there. I do not think we are more targeting with this for a medium-sized employer than may contract out to industrial hygienists or things like that. A lot of large employers are already doing similar things. Or they have their own mechanisms set internally to derive OELs. We are targeting folks that are not doing this now, but we would like to make it so that your mom and pop shop could do this. I do not think we are there yet.

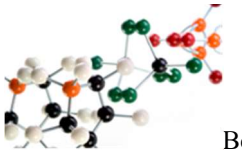
(16) **Question/comment:** If you have the data sufficiency score and one of them was already at E, which is the most urgent category, is the job not done?

***Response by Dr. Whittaker**:* We had lots of debates. I think you still want to characterize all of the health endpoints because if you had LD50 that put you in the E category or in this case skin irritation into E and that is the only data you look at, you may miss if it is carcinogen or something else. I think that is important information to have. It's tempting to say we just throw it into E, but I think you miss a lot especially because although we are most concerned with inhalation, we do have skin irritation. That tells us the chemical interacts with the skin. So, there is more power in looking at the whole toxicity and it will give you a better idea of what the hazards really are.

(17) **Question/comment:** Is there a way, in addition to the OEB, to explicitly characterize the level of uncertainty that goes along with that? You could say it's B 2, which means it's a solid B.

***Response by Dr. Whittaker**:* It does not give you much confidence in where you end up because you do not have a real measure of potency. For some of the longer-term studies, like if you have inhalation unit risk or something you might be able to do that sort of thing. I think some of the endpoints would be more amenable to this than others.

(18) **Question/comment:** Obviously, the OEB is all built around inhalation. You may have a compound is corrosive and you could have a notation, for example, for dermal exposure. A situation may arise where you could band in the E category for inhalation, but you get a drop or two on your skin and you are dead. Those instances have happened. Current banding can alert you to dermal exposure, but it is not called out specifically.



***Response by Dr. Whittaker**:* The key message obviously from the whole banding process is to make sure that you encompass the whole potential range of exposure. NIOSH is now looking a little beyond notation and to actually characterize the dermal hazard as well. We are concerned about dermal exposure as this is the other major exposure route for workers. This is one of the things we are trying to figure out how to operationalize.

(19) **Question/comment:** Comment. I wondered if physicochemical properties can be moved higher up into consideration and not just rely only on globally harmonized system (GHS) labeling. For example, if there is a skin or eye problem, we are always looking to make sure that we've got inhalation here while it dries. Thus, if there is skin or eye problem, you need to consider inhalation. You need to capture if there is skin and eye problem. If you need to decrease complexity, physicochemical properties and use profile should be counted as the first stop.

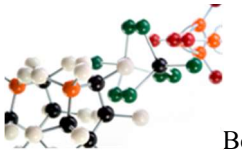
***Response by Dr. Whittaker**:* I actually think it is a hierarchical approach and I think of this in terms of different lines of evidence. We are trying to reduce complexity and not to increase. Obviously, those are kind of an obvious first stop.

(20) **Question/comment:** It looks like the OEB process is geared toward a single chemical, and I wonder if you are thinking of applying this to chemical mixtures or co-exposures.

***Response by Dr. Whittaker**:* We have not looked at that specifically, but that is a direction we can definitely look at, but it is important to look at. We could take this back to the cumulative risk to ask if there is a way to use banding for combined exposure.

(21) **Audience**. There was a brief discussion on inhalation TTC, but there are several publications on that now. Many of these publications do not include physicochemical properties. I was wondering if any or someone else has a recommendation on inhalation TTC approach that you would feel comfortable with and recommend.

***Response from a panelist**:* Unfortunately, my answer would be if it does not consider the physicochemical property, it is wrong. I am not aware of any to recommend, but you could make those available with this presentation.



This page left intentionally blank.

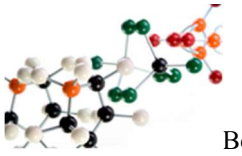
Case study 2: Use and Application of Real-Time Exposure Monitoring

This case study was presented by John E. Snawder, PhD, DABT, of NIOSH. The presentation is available on the *ARA* website. The purpose of this case study is to explore how to integrate and utilize the data from real time instruments into exposure/risk assessment.

Discussion and Comments from Panelists and Audience

(1) **Question/comment:** I commend Dr. Snawder for putting this together and wish it were available about two years ago. I would like to leave it to the epidemiologists to look at the methodologies/techniques and how you are applying some of these to the standard setting, as there are limitations to interpreting the data generated. I like the idea that monitoring can be done every second. The concept of using peaks and averages represents traditions that evolved from the best that we could monitor historically. Averages were being calculated before probabilistic tools became available and the desire to use one number that was not the minimum or maximum. This was designed for point estimate assessments. It looks like we are past use of averages and peaks and our capacity for data transcends the use of averages. The suggestion is to use the capacity that you have to help the toxicologist think anew about the metrics that go into the algorithms for hazard assessment. The tools you have with a method that has validation and confirmation can be used to assess profiles of exposure toward different necessary metric that relate to certain toxicology endpoints. For example, you could take an individual and monitor their benzene that they accrue from whatever tasks they are doing across their career in order to measure some type of total accrued dose because the person might be susceptible to cancer. This type of data could also redirect the toxicologist point of view toward looking at the safety window. This would provide information on what you anticipated to cause a health effect versus some type of exposure value. The upper part of that window was determined from toxicology studies where the animal was a virgin to the exposure, but workers are not. Workers may exceed a number/dose once in a while but if you have accrued tiny bits of exposure, what was the clearance between the last time they saw the exposure? Is it all out of your system or are you accruing in real time a dose even though your exposure might have been below a number you are worried about? One step beyond that is if you are working in an area where each exposure may be doing a tiny bit of damage but nothing that will be significant. There may be minor collapsing of cells in the peripheral parts of the lungs or you may have spots where your kidneys are not functional for the tiny damages but they accrue because exposure is ongoing and you cannot heal them and they do not get fixed with time. They will not be something you could pick up on one exposure but if you refocus the questions, you have an opportunity especially in a controlled situation like this to put together exposure profiles for real workers to take a new look at the metric we should be looking at in a temporal relationship or be asking of how we measure the risk and how to express dose in in terms of what data we should be collecting. One can also use this information for ranking an individual for probable susceptibility. For example, you may start getting nervous about continuing to send a worker who has been working years on those tanks because their window might be closing.

(2) **Question/comment:** It occurred to me that you have sensors that report instantaneous results. This seems to be an extension of what Dr. Whittaker was talking about where you go from a chemical to a smart sensor, making it to the toxicology of the compound. Some therapeutic effects



are seen next year, and some are driven by concentration x time (C x t). There is an opportunity with the power we have and more than enough horsepower to set up the right order written algorithm linked to the biology. So, the alarm just does not go off and the workers get stuck. The alarms only go off when they need to go off, coupled with knowledge ultimately built into the system of the biology and talks a logical response. Can learn from or adopt useful strategies from the automobile industry that use sensor technologies to make it more meaningful or functional? Cars are loaded with sensors and have sophisticated systems for when to go off. Sensing elements are available but lack the "smart" attribute. We need to link to toxicology characteristics of chemical to know if signal observed is important. Once the sensor is saying we know C x T is driving it so when does C x T cross a threshold? Not just when it goes off a given concentration, but it has to be integrated into an algorithm correlated with knowledge of the toxicology. I know that is a challenging task, but I suspect with today's technology, that is easily within the realm of doable. Postdocs are doing a lot of groundwork and developing C x Ts empirically and developing apps to do that.

(3) **Question/comment:** Sensors have greatest value as a training device. Just imagine knowing that you walked 5000 steps today. Sensors help change behavior and this is where we need to be going into the future. Helping workers understand the consequences of their behaviors can lead to change in behavior. If the sensors can be linked to an app on the phone, the workers can change their behavior.

(4) **Question/comment:** This is a very nice write up. To echo something that has already been said, I would suggest that the best way to evaluate sensors is PBPK. It is a good way to test your dose metrics. It is not just whether there is a peak but how frequent is the peak and the magnitude of the peak. It is important to redirect the derivation. From toxicology point of view, everything on your assessment worksheet is exposure driven and not toxicity driven. I think the way to turn it around is to say what the dose metric or biology of the chemical is that will result in toxicity. I have advocated that we should not necessarily be doing 6 hours/day, 5 days/week in vivo testing. Some should be a very different experimental regimen to understand the right toxicity. The right metric might provide the opportunity to turn this around. Volatile organic compounds (VOCs) have postexposure leaching. The exposure ends but the chemicals leaches out from storage tissues and you are still being exposed. It is not only clearance but the frequency of the exposure and magnitude and the best way to integrate that is PBPK modeling and they are relatively inexpensive.

Response by Dr. Snawder: I am writing a paper with Bob Hoover on this to think more about dosimetry and how it needs to be part of the conversation in the development and applications for sensors. There is a need to consider exposures, particularly with respect to time, concentration and how that chemical behaves in the body. There may be more personal exposures than we realize. That is the problem that could have potentially bad consequences. The actual toxicity of the chemical should be considered during the development of the sensors.

(5) **Question/comment:** I am fascinated by the feedback mechanism to have workers change their behavior. I am curious to the extent to which it has been studied or documented and the extent to which it is being resourced at the beginning of the process. It is really a powerful tool.

***Response by Dr. Snawder**:* Researchers in the NIOSH Mining Research Division are doing that (assessing worker behavior changes) a lot and NIOSH has started a joint collaboration with TNO in the Netherlands and HSE in the United Kingdom that is looking at different ways to assess exposures, especially cumulative exposures. We pretty much incorporated that (assessing worker behavior changes) in our first meeting. We wanted to change worker behavior and figure out how to do that. Video exposure monitor is a powerful tool to do that, especially if you can image in real-time data. A researcher at "Mines" has been doing work with the video exposure monitoring where she had one video she shared with the worker who did a great job and was changing out filters, did everything right and gave her a thumps-up. I am amazed that people say you can do that between seconds to less than a minute. I'm hoping to learn from that.

(6) **Question/comment:** That is the point. Had you gotten these key messages out in very short snippets? It is engaging the workers in that discussion. What would catch their attention? Clearly you have data that indicates that videos catch their attention. I really think that increasing impact in this area should be resourced well.

(7) **Question/comment:** We are sort of going past the area under the curve (AUC) versus maximum concentration (Cmax) and thinking about severity. So, trying to pick about how the severity part comes into the calculation. A very high peak could be fundamentally different than what one is concerned about over 8 hours. And that could be more than 100 times more severe. Most of the time, we figure it out the hard way. We do find that sometimes one has to look in the medical literature or look way back in the occupational health and safety literature. These are the places you can find the data. And the information you find is that they have intoxication. How do we take this toxicity endpoint at the high peak into account? Is one of the extensions of Jim smart from-based tool going to somehow include severity?

Response by Dr. Snawder: I hope so. Some people do severity based on concentration numbers and multiples of the occupational exposure limit and that is not necessarily what you want. It is good to guide your respiratory protection logic but that is it.

(8) **Question/comment:** The trade-off we're talking about but not really talking about is how many false alarms you will get from this thing. I would tolerate an awful lot of false alarms for something that will put me in my grave. If it is going to be irritating, not so much. Just wondering about that logic and the severity.

Following up on Bette's point and your point about the behavior change of the workers. This can be wrapped into the risk assessment concept. You are talking about how to get the right alarm at the right time. But, if you can create the conditions where the situation does not happen in the first place because they have appropriately taken care of what the workplace danger is, then you are reducing the risk and the alarm does not have to go off at all. It is almost like a right sensor right time and then something else. You can work on the acronym later. Is one of the extensions of Jim smart from-based tool going to somehow include severity?

Response by Dr. Snawder: We worry a lot about alarm fatigue. You can buy a monitor that

comes from the factory and the low alarm for carbon monoxide, which is set to 35. There are a lot of jobs where you walk behind trucks or vehicles and we recommend people set the low alarm at 50 or 100 ppm and keep their time weighted average alarm. That way, you knew you used it in the wrong place too long. Alarm fatigue is tough. We do run into people who do not get worse because it goes off all the time. It is going off when it should. We need to adjust the time a little bit.

(9) **Question/comment:** We live north of the motor city and we love our cars and car technology. My husband's car has different sensors on it. If I do not use my blinkers when changing lanes, I can feel it and I cannot do it but that is a subtle sensor for me to change my behavior, not telling me there is any danger. I think Jam is onto something about the automobile industry changing our behavior with different types of licenses and vibrations. Then, there is the dramatic avoid a collision. I think that is a cool analogy.

Response from Dr. Snawder*:* In our motor vehicle program they do a lot of monitoring and fatigue sensors. Those are the ones where you do see behavioral changes. Because they will vibrate or whatever. Or the tingle or one that looks at brainwave seems a little creepy. I agree, the more we can do that the better.

(10) **Question/comment:** I learned to appreciate the new models where they have the lighting sensor. If someone is in my blind zone, it will flash at me and it reminds me, because occasionally I find myself lazy and not looking over my shoulder and that light will remind me. It is the feedback centers that could be applied as Carol indicated.

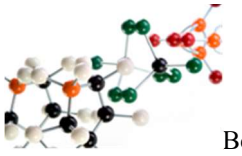
Response by Dr. Snawder*:* The amazing thing is if you have a rental car today that does not have it, how much you miss it. Even backup cameras. I do not know how I use my mirrors to back up.

(11) **Question/comment:** Great presentation: Speaking on the severity thing, the EPA has the acute exposure guideline limits (AEGLs) of 1, 2, and 3. Can you use something like that to assess peak exposures?

Response by Dr. Snawder: That is in our back up to. AEGLs are one thing along with the NIOSH Pocket Guide.

(12) **Observer**: I have a question. I think what I have heard so far is the use and data interpretation and application of acute exposure. Observer: What we heard so far is the use and data interpretation and application of acute exposure. When we have this technology, how do you measure repeated exposure? We have a scenario where we have monitoring results. The exposure generated is once per week for 24 hours. Do you do the mathematical average and average up to one year and consider that as chronic exposure? If we have this from the duration perspective and data perspective, then how will you use this?

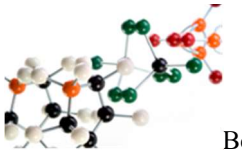
Response by Dr. Snawder*:* That is the thing we are running into more. Hopefully, next year on a NIOSH National Occupational Research Agenda (NORA) Intervention and Control project, I



will subscribe and get 10-real time monitors and cast our fates to the wind. I will get one year's worth of data on up to 10 workers for a given year. It will be able to integrate that. I have a selection of sensors I can put in. All day, they are uploading to the cloud and I get the dashboard and can look at the data. Our next step is to start doing this in the realm of the data. We identify the task and then start on how to better integrate them. Dr. Abbas Virgi of NIOSH has been developing statistical tools to do this. So far it is based on healthcare workers. He has been working on models to analyze or manipulate the data and do postdoc analysis and integrate that into the time weighted average but not ignoring the peak. I hope to do something like that in my next task, as soon as I get really smart people to help me handle the large volumes of data. I am taking baby steps right now.

(13) **Question/comment:** Are any of the sensors being developed to look at co-exposure or simultaneous exposures to the endpoints that have the same effect? I'm talking about ones with the same toxic effect. Is that on the consideration list?

Response by Dr. Snawder*:* It is on the consideration list. Sometimes, we see if we can leverage cross sensitivity to our benefit. Alkylamines/organic amines are really nasty and we're finding them more and more in my world. They are very hard to measure but not hard to find in use. We may not have the ideal or most accurate way to detect and measure but we may have multiple methods that are ok or good enough and used together in real-time could be far more protective than the perfect laboratory method that takes days to analyze.



Case study 3: Applying Hypothesis-Testing methods to Help Inform Causality Conclusions from Epidemiology Studies

This case study was presented by two toxicologists from the Texas Commission on Environmental Quality (TCEQ): Sabine Lange, PhD, DABT and Lalita Shrestha, PhD. The presentation is available on the *ARA* website. The purpose of this case study is to develop tool(s) to better understand and interpret epidemiology studies in a way that informs the causal conclusions that can be made from those studies, bearing in mind the limitations of observational studies under consideration.

Discussion and Comments from Panelists and Audience

The discussion addressed the five major considerations presented in the preliminary approach.

#1. Theoretical basis for exposure variability

(1) **Question/comment:** In some of your graphs you show data. There are more data points in some of the later graphs than the earlier ones and that is what confused me. I would like to know how the simulated data sets were created. You are also fitting the regressions to simulated data that showing the difference in the slope. Is that how it works? You also went through the different types of exposure error well but at one point you also threw in misclassification. It wasn't clear to me that misclassification actually applied in the regression. And in what you presented, you did not mention it, so you need to clarify that.

***Response by Dr. Lange**:* The simulated data were all made up. The regression lines are just to have a visual representation to show the concept/pattern that we are trying to demonstrate. You are right about the misclassification - the concept of misclassifications is not addressed in this case study. I think partially because it had become clear that this probably wasn't something worth pursuing further, so probably not worth going down the misclassification path.

(2) **Question/comment:** One thing I would like to suggest is working with epidemiologists to find out if one can quantify the amount of variability or misclassification. Or with the data you have, can you quantify the amount of variability or misclassification or both? The objective is to more systematically quantify the variability of both exposure and outcome.

***Response by Dr. Lange**:* Yes, that is something to consider. I think that even though it is something we did not pursue past this point for this particular study, it still remains a very important concern. One of the reasons I included this is that people make this assumption, and we want to have a better idea about how much uncertainty in the exposure error in classification is leading into an estimate.

(3) **Question/comment:** It just gets to the sensitivity analysis that if you can estimate and quantify the variability or misclassification and you can include it into your model and determine what if any impact it has on the ultimate risk estimate.

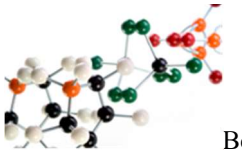
(4) **Question/comment:** I really like this case study and just on the point you are raising here, I would not have continued to pursue this just based on my experience in considering epidemiological data. I think the message back to the epidemiological community is that when we're quantitatively estimating risk based on epidemiological data, we need the original exposure data. For risk estimation, characterizing the variability in exposure is critical and we rarely have sufficient of this information.

(5) **Question/comment:** I am trying to figure out from what you said whether these conclusions are sort of conditional upon assumption of a linear relationship, because all the information you presented was what happens to the slope of a linear regression. Another related question is, if you are pretty sure there is some sort of a non-linear relationship, maybe not so far as a threshold but something with more significant shape to it, would you not expect the misclassification or exposure variability to be enormous? You used multiple monitors in the city and used the average of them. Surely you cannot ignore that.

***Response by Dr. Lange**:* Yes, I would not suggest ignoring it. I don't think, for the purposes of what we're looking at here, we are able to say that with better exposure estimates we would expect higher effect estimates. Those will not be predictable based on increasing error or decreasing the error. Using one of the case studies that we have, we're looking at lung cancer in smoking, we have the exposure estimate from surveys. Will it be much more reliable, much more precise than mailing questionnaires, so can we predict we would have a higher effect estimate? Is there a predictable effect of that? I think the answer is, no. I do completely agree that it is 100% important in terms of what to do with exposure error to certainly acknowledge it, to try to quantify it, to try to present it to a risk manager while saying, you know this is a big kind of unknown in the estimate. We looked at linear and nonlinear in terms of the relationship because mostly in my experience, those are what are used in the epidemiology studies that I was looking at. Most of them were applying regression that when it came down to the linear regression or nonlinear, and the effects I presented apply to both linear and log-linear regression. Certainly, we are interested in pursuing what happens if you had a threshold into the mix. These simulation studies are just a start, based on what I have seen as common modeling choices.

(6) **Question/comment:** This might be a bit related to what has been said already, but experience seems to be that as you operate in the higher dose range, confidence around the exposure is greater than often in lower dose ranges. So, current approach shows the variability to be constant, but in fact, I expect variability gets substantially wider as you get into the lower dose ranges from, sometimes, just analytical determination of exposure, whether that's actual measurement by some technique you were using, or by surveying or whatever versus actually going out and measuring it. It would seem to me that that enhanced variability is, again, in the lower dose range. Obviously, it would seem to be to change the slope again or biased towards the null again.

(7) **Question/comment:** I think the message is, it depends, because we saw the graphs yesterday of the workers where they had those peaks and a lot of times we will categorize our occupational exposures into the high peak if you can imagine you would have a worker or a dozen workers



that would be category two -- that is high but you might have an incredible range they had had that peak exposure like we saw yesterday, whereas the unexposed or low exposed might have been the office workers, and they would have very little variability. In contrast, we look at maybe some biomonitoring data, and that is the key, where the variability in the differences may be really great in the lower zone, so I think the message is buyer beware.

Response by Dr. Lange: Yes, it depends and, in this case, it's not going to work for us (as a pattern for causality determinations). Although, on a case-by-case basis and in a particular case, we may be able to predict the effects of exposure error. As a general rule, it's not going to be helpful for us. And even then, we might predict that a certain effect would be the case just based on the exposure error, but the problem is things change when you introduce extra coefficients into the regression. When you have extra coefficients, different exposure errors can impact the health effect estimate.

(8) **Question/comment:** Observations made earlier by a panelist is absolutely correct. I was thinking more along the lines of analytical variations, so say you talk about general population exposures, the office worker versus the plant worker by way of example. Your method maybe such that down at the office worker, you're operating around the limit of detection, so variability detection becomes a lot noisier than if you are at the OEL level of 100 parts per million by way of example. So, again, that would just substantially impact, or potentially could impact the slopes of those. And if you go higher than that then, again, variability can increase. So, you're actually looking at multi-faceted variability that is going to be problematic in these circumstances, I think. It's hard to anticipate.

(9) **Question/comment:** Complexity is so great that I am not sure at what point in time we would have confidence we would actually get reasonable measure of what you are looking at, which is the pattern across studies, etc. I think there are a number of other issues we could raise on variability and this rather limits its potential application, as considered here.

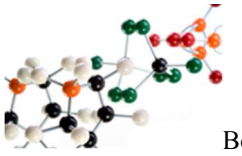
#2. Theoretical basis dose/exposure-response

(10) **Question/comment:** Regarding the concept of dose or exposure that you are using here, this is magnitude of delivered dose or whatever, the exposure, but if you are looking at something like response for cigarettes to cancer, where is the time element? So, for anyone spot on the magnitude a line? Okay, but that is combining dose at any one point in time.\

***Response by Dr. Lange**:* Yes.

(11) **Question/comment:** If you essentially made plots three-dimensional, for any one point on the magnitude line, you had an axis going into the plot, which would be time, wouldn't your variability fall out?

***Response by Dr. Lange**:* I'm not sure the variability would necessarily fall out. I think with time, you could get an increase in variability because of the variability in exposure over time, but



it's certainly a reasonable point to say that time does contribute too dose. It would impact kind of that, yeah.

(12) **Question/comment:** The reason I am asking that is if you do magnitude times time is a point, you really combine to do variabilities to begin with. And the other thing is you are making the assumption that low-dose, long time equals high-dose, short time. And I do not know where you have taken two factors and you have assumed that they are directly related. Is my understanding right?

(13) **Question/comment:** We get into trouble when we tried to graph it, but if you think about a multiple linear regression you can put time as its own variable. You can put intensity of exposure independently and compare those independently, or you can combine impact figures. Think about smoking, how many years have you smoked X, when did you smoke, how many packs did you smoke. You can fill it with 20 variables. Dr. Shrestha was just graphing one, but the key is, and this would this would be my message this afternoon, is to epidemiologists to not just stop your analysis with the regression coefficient and say that you are done, because TCEQ wants to know, is your assumption of linearity valid? Can you categorize or show the shape of the curve rather than stopping with that complex multiple regression? You are not wrong that time and intensity have their own variability, and so, epidemiologists can do them separately or collectively in their complex analysis, but does that make sense?

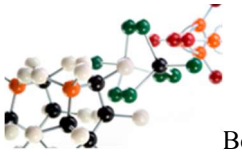
***Response by Dr. Lange**:* You can fill it with 20 variables. Dr. Shrestha was just graphing one, but the key is, and this would this would be my message this afternoon, is to epidemiologists to not just stop your analysis with the regression coefficient and say that you are done, because TCEQ wants to know, is your assumption of linearity valid? Can you categorize or show the shape of the curve rather than stopping with that complex multiple regression? You are not wrong that time and intensity have their own variability, and so, epidemiologists can do them separately or collectively in their complex analysis but does that make sense?

(14) **Question/comment:** Yes, it makes sense but how do you test whether the magnitude times time is always the same? In other words, is the effect the same if you had short small time, big magnitude?

***Response by Dr. Lange**:* Normally, we keep them separate.

(15) **Question/comment:** If I we're doing the analysis, might combine different kinds of acute exposures and consider that as a dose-response but would not necessarily put all durations into the same analysis, like the effect of one hour exposure with the effect of lifetime exposure. Like I said, there is a certain kind of relationship there. There are some equations for to extrapolate that but it is an entirely challenging problem to integrate the effects of time in short term and long-term. I try to keep those separate. Not something we tested necessarily, but when assessing the study, I would be reluctant to combine that much variability in a dose estimate.

(16) **Question/comment:** This really applies to severity but gets to your issue of the dose-response basis. The disease dimension is not being addressed. For example, in a short-term exposure, you



would not necessarily see cancer because the disease hasn't developed yet, so that is the other aspect of this where I felt like the graphs were not too informative. Because if you are only looking at one effect, that is not really what is happening. Disease outcome is just as complex as your previous analysis of exposure. The disease is not the same at anyone point in time either and that needs to be to be considered. So, or you looking at a precursor key event? Are you looking at an adverse outcome? If this is meant to just be only in the end-stage adverse outcome, that is one analysis. If your trying to capture the complexity of the pathogenesis and the disease dimension, I am not sure this is the appropriate approach.

***Response by Dr. Lange**:* A panelist and I were talking about this earlier in the week. We don't have any set outcomes yet that we think would apply to only the most severe outcomes, the ones that we used for the case study were cancer. But I think there is certainly an argument to be made for looking at the earlier effects and how do you go about doing that is something I am very interested in and discussing with people about how that can be applied.

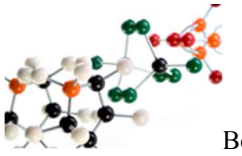
(17) **Question/comment:** I'm saying this from a mechanistic toxicologist point of view. It's a problem to try use epidemiological data to interrogate some of these mechanistic, biologically plausible aspects. You need other types of data to help inform this.

***Response by Dr. Lange**:* Yes, and certainly not something in isolation but rather to interpret the type of data. I think there are more things we could look into, for sure.

(18) **Question/comment:** Yes, along the same lines I actually think that I see this in looking at empirical evidence for epidemiological studies. For considering the empirical evidence from epidemiological studies, the parameters you are looking at are useful hypotheses, and I think many of them are exactly what you need. Beyond that it is coupling the empirical evidence with biological plausibility. This is akin to what we do in looking at empirical support in adverse outcome pathways (AOP). The relevant patterns also need to be considered within the biological construct of what constitutes a precursor event and adverse outcome? This is relevant to the context of how you are looking at severity of effect. Empirical support for the associations in adverse outcome pathways is based in part on consideration of particular patterns of severity based on the biological construct. But I do think there is a way to marry these together really nicely, and that is consistent with what we do in AOP /MOA. What we think we miss often is kind of looking at these patterns across studies and levels of organization and doing it with some kind of, as you say, less subjective measure, rather than looking at individual studies and saying, how do they contribute in the weight-of-evidence context? I really like what's being proposed. I think it's just a question of putting it together in biological context. We can talk about that maybe at the end.

***Response by Dr. Lange**:* Yes.

#3. Theoretical basis for specificity



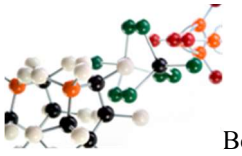
(19) **Question/comment:** This depends, in my view, about what you know about the pathogenesis of the causal association with disease. Is a more specific outcome likely to be more informative? If we consider the lung cancer example, if you look only at "all cancers" you will diminish the signal because not all cancers (other than lung) expected to be associated with smoking. Also, relative to pathogenesis, are we expecting to see specific types of lung cancer associated with smoking? The other aspect here relates to relative quality and extent of the more specific vs. nonspecific data which need to be accounted for interpretation. If you have 10 times as much nonspecific data "e.g., lung cancer" versus specific forms of lung cancer which might more likely be associated with the causative agent, this will need to be weighted. It depends, also on relative sample sizes, etc. There should also be some background information on the likely ratios of specific to more generic, non-specific effects.

(20) **Question/comment:** One thing I will not spend time on – it's mostly about food for thought, is whether or not that relationship can be used. The fact that you get a predictable effect of nonspecific data to specific data, is that going to be used to predict what a nonspecific relative risk would be or what total risk would be based on the specific? Where can we use it quantitatively to make a prediction about what the metric would be? This is just a thought experiment at this point and not something to be pursued. It is something that would get complicated really fast. That just as I thought maybe going forward if that particular relationship to be a bit more quantitatively to get at this specificity point. It was just a question in relation to, what would you be wanting to use it? Again, if you're looking at patterns across the data, it is in the case where you had a lot of nonspecific data that could contribute to your kind of characterizing the pattern for the more specific data. Is that how you would like to use this?

(21) **Question/comment:** It's more about a little bit, so, the pattern would be that if you have smoking in lung cancer, you would expect to have a higher effect estimate than the total cancer, but how much higher? And can we use that information? If it's two-fold, does that make sense? If we have 10 times as many non-lung cancer cases, but lung cancer only two times higher, does that make sense in terms of what we think the relative risk is? Just an idea if we can get a little bit higher than, lower than, can we get it to further saying how much? I am thinking you would rely on the more specific endpoint in any case, but in terms of characterizing the pattern, I understand, right. It seems attractive enough but may not be accurate so that's what I'm putting out there. It may just be to simple. I do not think it's necessary, but I think it's a new hypothesis-testing exercise.

(22) **Question/comment:** I had a similar question to previous panelist's, and then also I am wondering if this isn't part of what we said earlier about dose-response in maybe getting to the severity that we have not discussed yet. Picking on specificity, it seems that this did not take into account the prevalence, or if it is a rare occurrence. I get what you're trying to do in terms of signal to noise, but I was concerned about how you factor in the actual prevalence or occurrence of what you are trying to see to begin with.

***Response by Dr. Lange**:* Right, this is a little bit about trying to get at that. Maybe there is a better way to do that.



(23) **Question/comment:** What I hear you saying is you would like to be able to use the nonspecific outcome data, based on what we know about its relationship with the more specific. My message earlier this afternoon was to better communicate with epidemiologists about what we need to inform risk assessment. To me, this is going down a murky path of accepting and being okay with lesser data, rather than communicating with epidemiologists regarding priorities in data generation and reporting.

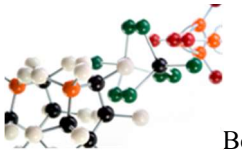
***Response from Dr. Lange**:* That is a great point and I like that idea of communicating what we need, better trying to make sure we have those conversations. But this method helps us to assess work that has already been done. That is part of what we are taking into consideration.

(24) **Question/comment:** One thing I am trying to tease or guess out of your conclusion is that it seems to me you are using the attenuation of the slope as what you are talking about. But, in reality, what this is ultimately leading to is determination of causality, which is really a question of the slope. The question of whether you think the slope is real or is that causing effect. Even if you know that the specificity is going to affect the slope, the question might be extended to what is the real threat to your ability to make the causality determination because of the lack of specificity? In other words, there is a real effect potentially hiding in a cloud of nonspecific data. And then as a secondary question, but along the same lines, would this apply specifically to a subpopulation that is truly at-risk hidden within a group of population that is otherwise not at-risk? For example, the real effect is among asthmatics, but you got a population which includes both. Is that conceptually similar to the specificity issue here, because I did not see it in specific issue in your list of things to think about?

***Response by Dr. Lange**:* I think it would. I would have to think about it a bit more, but I think in theory it's the same signal-to-noise problem. I like the idea of that a lot. I t's harder too, a lot. Again, looking for patterns, you have to have somebody who did the analysis first to be able to test to see if there was a pattern there, and this we can test to see if there is a continuation of the effect if someone had tested both for lung cancer or whatever. Somebody would have had to assess a whole population and then just people with asthma within that population for us to be able to see if that then applied. think the idea is good, but I think you have to have the data to apply it.

(25) **Question/comment:** Right. I am using the slope as the indicator but it's being used kind of as an indicator of if this is a real relationship then there should be these signs. We cannot get directly to the causality question, at least not with most of these data, if we are thinking of epidemiologic data. There should be patterns that exist in that epidemiological data. The attenuation of the slope is an easy way to look at patterns. In fact, often it is the only outcome that is being provided to us, the only measure of relationship that is being provided to us, so we kind of have to focus on that measure relationship, but I do appreciate what we have with causality.

(26) **Question/comment:** Yes, we're not looking at estimating risk, here, but rather, patterns across identified epidemiological studies for various aspects of causality (e.g., the Bradford Hill considerations of temporality, specificity, dose-response etc.) We're looking at those patterns to try to get a handle on what elements of causation the collective data fulfill. This is based on



measures of hazard or risk for each of the individual studies. For quantitative measures of hazard or risk, we focus more specifically on better quality studies, where we have most confidence of assessment of exposure and hazard. For the cross study considerations related to causality, we are looking for indications of patterns in the data, that inform us maximally on support for causality. So, the question you are posing is where we have the less specific effect, are we able to characterize uncertainty based on its relationship with the more specific effect? And that would be very valuable information, because we don't always have all of the data that we need.

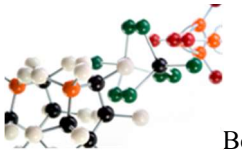
#4. Theoretical basis for severity

(27) **Question/comment:** Some of these criteria are similar to Bradford Hill's. I've been struggling with epidemiological studies for my whole career, and just to take an example of particulates and mortality. I keep thinking to myself, why should we plot mortality in a linear fashion? I'm directing this to my epidemiological colleagues. When I know there's less severe effects going on and they should be greater, and we should be able to measure them, and we really should focus on that part of the adverse outcome pathway, or whatever word you want to use, because we're going to get a better does response assessment. But anyway, to this case study I finally saw we could probably do something along these lines. I really like this a lot. The only caution, and it's not much of a caution, is even in toxicology studies we often do not plot the different severities of effect in animals. We often will look at the most severe effect, and we account for that. And so, what you get is as you get higher and higher exposures, there is this drop off of the less severe effect, not because they are not there. We did not count them. So, with that, you know, like washout, I think this is something that I really like, and of course, I would like to hear other colleagues' comment, of course, on this.

(28) **Question/comment:** What's being proposed here is absolutely consistent with how we assess empirical support for AOPs, though we're using different terminology. These are well known and accepted patterns of empirical support, based on biology including that for what is being described here as the "theoretical basis for specificity". This are the kinds of relationships you anticipate if empirical data support a hypothesized biological description of disease pathogenesis. So, I think it's a complementary corollary to what we are doing on adverse outcome pathways and we probably need to start using similar terminology to promote common understanding. The pattern analysis being considered here is completely consistent with our experience in considering MOA or adverse outcomes. I think there are ways we can integrate biological plausibility here as well; we can talk about that later..

***Response by Dr. Lange**:* Okay, it would be really great to integrate with some of the biological pathways.

(29) **Question/comment:** I also like this idea of plotting it this way but, suppose you had a situation where you ask a question as whether what I'm going to describe would apply to this. Suppose you have a situation in a town where everybody at a given moment in time suddenly has a contaminant in their water supply, and the contaminant delivery is the same across community. So, you got equal time and you got, theoretically, at least equal concentration, which doesn't necessarily mean equal dose, but it certainly takes a lot of variability out, certainly, variability in



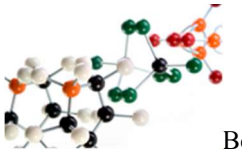
smoking or something. And then you look at the community to assess potential outcome. Let's say for this example that it causes freckles and you can count of them. So, even if in this hypothetical case you are removing some of the problems, would you expect, in order to comply with your hypothesis, that you would have to have the same slope? Or could this be a different susceptibility issue within the population, breaking out to more severe, less severe, median severe, not because of dose effect but because of susceptibility in the population? Could you use this basically to look for that kind of thing?

***Response by Dr. Lange**:* This specific case study doesn't address whether or not the difference in severity is because of the difference is susceptibility in the population. I think you would expect to see some people that don't have any freckles and some people that do have freckles, some people that do have a complete pigmentation change in melanoma or something like that, a more severe effect. I do not know that we know it's in an eight their ability in a [Indiscernible] kind of way or some people like myself predisposed to getting freckles. Doesn't have that difference? I don't think it is restricted because of why you would have the difference in severity.

(30) **D Question/comment:** This conversation just illustrates the chasm between epidemiology and toxicology. Dr. Sabine Lange already knew where I was coming from. I do not like this at all. Sorry. And we just have to look at the cruise ships in the with people with coronavirus, and it's about everything is confounded of severity. It's about susceptibility as a starting point. I t's about stage of diagnosis. Every woman in this room should be getting mammograms, and so, we should not be adding mortality from breast cancer because we should be having early diagnosis and early cancer treatment. Which is confounded by our motivation and our healthcare and our ability to comply. So, by the time the epidemiologists are getting around to doing a case control study, this control study, this is washed away. You have to control for all of the other things. And so, you can't tell severity in your epidemiology studies. So, there is an example that, may be, would argue for this. Something like sperm count or IQ that is on a continuum and you can't treat it, but then I would argue that maybe instead of severity, is diagnosis outcome actually adverse? Are you measuring IQ points? Sperm count? Cholesterol points? But there is a point at which that outcome adverse cost sort of more about binary than the severity, because that will build into it but there is, you know, a whole discussion in the chemical literature of top is A1 point IQ, is that within the variability? Is that really adverse? I am thinking of severity from the environmental chemical, like, to say once I do a study with asthma attacks and then mortality and then concentration. Is that all in a way? You can have a new expectation of a slope there, do you think? The asthmas are going to be the severity of the asthma. Is it yes or no? That's not severity. So, keep it within an endpoint.

***Response by Dr. Lange**:* That's the asthma and hospital admissions and that's a more serious. Did you think that's also something that could not be expected in this focus because of the confounding?

(31) **Question/comment:** That's not going to give you the information about causality that you're looking for. It's driven by personal behavior as well as severity. You're talking about repeated. I think it's a slippery slope. Cancer is not that example. Maybe something more like asthma.



***Response by Dr. Lange**:* Yes, I agree. We pulled that one out too. I agree.

(32) **Question/comment:** Even with asthma, it goes to what is the diagnosis? Are you exacerbating? Is it a chemical that's triggering an attack, or is it the chemical that caused the attack? Two different questions.

***Response by Dr. Lange**:* Yeah, and then exacerbation gets back to those.

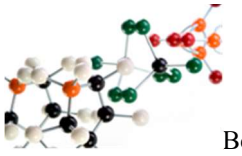
(33) **Question/comment:** Yes, that's a time thing. For asthma, it's a key and long term and stuff like that, and I agree that for that one it is problematic.

(34) **Question/comment:** I have a comment on the example. When we look at severity, oftentimes the higher concentrations are masking precursor events, so that those events are no longer observable. Factoring in disease dimension is needed for causality - are early events actually causal?

(35) **Question/comment:** Greg's point is important for looking at the causality. You need to assess what the probability or the confidence that the key events earlier are going to cause to the adverse outcome. I want to point out that this is disease dimension. It's something that for example the America society uses to characterize biomarkers. And with respect to categorizing biomarkers, we have Dr. Schulte the audience who developed mathematical criteria for biomarker validation. So, I think there are ways to address this supposed continuum. You have to have precursor events that everyone agrees are related. That gets back to the biological plausibility and AOP context. I don't think that you can do it with a single measure well. The entire set of measures characterizes the disease. What are the surrogates that you're using along the disease dimension? That gets to Chris's concern on temporal aspects as well.Is there a pattern of disease that we clearly agree upon? To the point that they dislike this is because they're looking at the wider spectrum. Can we get into this intermediate layer where toxicology and epidemiology intersect? I do agree that you have to lay it out. What are these? It's not just kind of effects that have been multiple responses in the epidemiological studies and making the determination that one is more severe than the other. There are ways in the biological context to address all of these confounders. They are called different things. They're called modulating factors. We have agreement on a number of disease pathways now. I think that's a good starting point, actually. I think that some of them have moved into the biological monitoring realm. This requires the interface that we have been facing in epidemiology and toxicologist prior to doing the studies.

(36) **Question/comment:** Yeah, that's the best time to catch it. The pattern is clear. The words maybe concerning to some and we have had difficulty of dealing with this in relation to AOP because different communities interpret the context of "severity" very differently.

<u>#5. Apply theory and practice – positive and negative controls</u>



(37) **Question/comment:** Thank you for that clear example. The question that I had is whether the epidemiology data allow you to make a distinction between cancer and something less severe. I don't think COPD is along the AOP to cancer. Did the data allow you test that hypothesis?

***Response by Dr. Shrestha**:* I don't think so. At least for these studies, we are not looking in terms of the adverse pathway, but we're looking for specificity outcome.

***Response by Dr. Lange**:* We were keeping an eye out for that as we were looking for the studies. We did not see a lot of evidence for that. That would be really interesting to see if there are precancerous lesions or something like that. We briefly toyed with the idea of looking at severity, incidence and mortality, but again with the confounding we are not quite there yet to being able to use that as an endpoint.

(38) **Question/comment:** It might be good idea to think of the source of the data. If you're talking about mortality, you are looking at competing causes of death. So, you are going to have death due to COPD, and thinking is that they are typically related to cancer. Because it's a mortality study, you cannot evaluate -- you only die once. You would have to look at a different design that's looking at progression. That's not what these studies were designed to do.

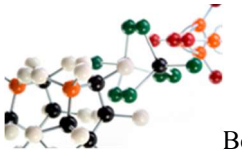
(39) **Observer**: For the estimates based on the six studies, the highest was really high at 40. I would be worried about how stable the intercepts are with the nonsmokers and relatively few lung cancers. Smoking is generally assumed to account for 90% of lung cancers in society. So, 40-fold relative risk is kind of high.

***Response by Dr. Lange**:* It's certainly something that we will look at when we further investigate those studies.

(40) **Question/comment:** Again, a great presentation with examples. The question I have is more practical. Now that you have been through this, TCEQ looks at lots of epidemiology studies. How are you particularly doing things differently?

***Response by Dr. Lange**:* We have not changed anything yet. We want to prove this theory and make sure that it is sound and then we will try to apply it. My thought is that if it comes out looking good and we establish it, publish it or work with other groups to solidify it, we will start looking for this type of data. We will try to apply these patterns, mostly the hazard demonstrated by epidemiology studies. It sounds to me, first of all it needs to be further developed, but also I think it would be useful to integrate more into what we heard earlier on from Bette with the AOPs and trying to pull a biological aspect of that in so we can put it all together. This is a great overview and a way to look at the epidemiology data. We want to use that also to integrate of course the animal data and any other information we are getting.

(41) **Question/comment:** Bette and others have published on these modified Bradford Hill criteria for risk assessment. If this is successful, do you see this further modification or enhancing what we already have? Maybe a question for both you and Bette.



Response by Dr. Lange: I think it is a compliment to that. I am not sure that it changes that or exactly how it would fit into the modified Bradford Hill criteria. I think at the very least, it is a complement in theory.

(42) **Question/comment:** I really like it. I think it extends to what we are doing on AOP and we could also contribute to what you were doing. A couple of points. Just one in relation to what has been raised in terms of how you might envision using this. So, when we do MOA and AOP analysis we look at the weight of evidence considerations in what I call a top-down approach. So, were looking at the Bradford Hill considerations and those that are most important, and they are modified or defined to suit purpose. We give examples of datasets associated with high, moderate, and low confidence. In the sense that you are doing a first cut from a top-down approach to determine, do you get consistency with the patterns that you expect initially, and then you could drill down additionally to understand why you do or don't have consistency with the patterns that you expect. I would expect the patterns that you have identified based on the experience we've had on AOP and MOA. I do think it requires some of the caveats that others have mentioned related to integration of information on the disease progression which needs to be factored in. I suspect that some of the associations, e.g., the specificities can be very difficult to address due to limitations of the studies. Also, there is always overlap, interplay between the different studies and patterns, so we always look at why you have low support for some patterns but high for others and why. All of this relates to cross study support for evidence of causality based on various of the Bradford Hill considerations and is entirely consistent with approaches for AOPs/MOA. I'm thinking that on this basis, biological plausibility could be additionally addressed. There is a very nice example where EFSA has looked at the plausibility of associations between pesticides and Parkinson's disease, based on a documented AOP. Consideration of biological plausibility based on an hypothesized AOP addresses the aspect of disease progression to consider the nature of relevant patterns, such as specificity. I would certainly like to follow up with you and see if there is any way we can integrate these approaches. I think they are completely consistent and would inform each other. I'm very pleased to see a more top down approach to looking across the epidemiology data; this is critical, in my view, to focusing on the aspects which impact the outcome of hazard and risk assessment to the greatest extent. In this context, we've been interfacing with the systematic review community, which is a very bottom up approach to assessment. The objective is to integrate more top down and bottom up approaches to focus efficiently on the more critical aspects which drive the outcome of hazard and risk assessments. In my view, these top-down approaches considering patterns across a database need to figure more prominently. I really appreciate the case study and commend you on thinking about it in this context. It is really progressive.

***Response by Dr. Lange**:* I would like to work more with you on it.

(43) **Question/comment:** I thought Bette would comment more on the AOP. I was thinking about the specificity of the outcome. If, instead of talking about cancer, you're thinking about respiratory outcomes, then COPD and lung cancer, are they together? And I guess the question would be, do you know if your chemical is acting and what that mode of action is? Is it a respiratory irritant or whatever or a carcinogen? And can you lump or divide? I think that is an opportunity to build on what is known about the chemicals and to incorporate that into the step to

build that into biological plausibility first, because in epidemiology, we do that last. Show us a better way.

(44) **Question/comment:** I will just reiterate again that I think not only the AOP context but also the biomarker validation context is especially useful because we address specificity and sensitivity of biomarkers in terms of categorizing them. Are they a measure of effect, exposure or susceptibility? I think you will find those papers informative in terms of looking at the disease dimension and the probability that it is causal.

***Response by Dr. Lange**:* Yes. That would be great. Thank you.

(45) **Question/comment:** This might be a basic question, but I am trying to go back to the AOP roots. When we look at AOP as a chemical agnostic pathway, we are looking at lung cancer and smoking. And so, to apply this for something else I would just try to connect the dots there going from a chemical agnostic language to making an AOP pathway and how we can use it for other examples. And the second part of the question is, is this only for data rich and can we use something similar for chemicals with a limited database which we often have or outcomes that are not that clear? Maybe Michael asked that question.

Response by Dr. Lange: I think we started out with examples that were very clear in terms of causality and it would be worth it to step back and see if we can get something that is a bit more refined in terms of chemical positive control kinds of examples or negative controls to see if these patterns still apply. I think the concept should be the same no matter what. If it's an epidemiology study and we're looking at association and we think it's a chemical causing affect, we would expect that chemical to show a dose response and have outcome specificity patterns. I would expect it could apply to these other scenarios. In terms of the database, it would have to be something that actually has epidemiology studies done on it which is already fitted in our toxicity data. If you don't have any epidemiological studies, this is obviously not going to apply to it. It is already past the LTD chemical stage and, really, we would like more than one study so you can look at patterns. It is a way of looking at patterns across different studies and study types in different directions. I think in theory this could apply to looking at patterns across studies and you would have to have more than one to apply it to. It is a more data rich method.

(46) **Question/comment:** I guess in a lot of studies we look at lung cancer for a specific chemical and smoking is a confounder. You are looking at the other way around.

Response by Dr. Lange: Right. And a point that I was trying to make too is we went through that and it's on our final slide. This doesn't excuse the need to look at the data quality and make sure the study was done appropriately and all of these confounders were done and a sample size and all that stuff and to interpret the results of the study at this pattern level. You still have that fundamental need to incorporate the appropriate confounders and study quality criteria.

(47) **Question/comment:** I had a comment on that too and if I could build off that and the data quality issue. It would be helpful, I think, if you had some of that upfront what the criteria were

before you pick these examples going forward. Even in what you are calling the negative control examples, if you have to study showing it equivocal, they were showing positive. Where are you going to draw the line? Are you picking examples to fit because it is showing the ultimate direction you want? I know you are not doing that but it gives the appearance if you don't have the criteria upfront about the data quality for individual studies and how many, otherwise it looks like you're picking to fit the examples that you want.

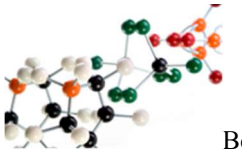
Response by Dr. Lange: Yes, and using the systematic review methodologies would be helpful there just for the transparency reasons that we use similar stuff to help identify studies we are looking at, to incorporate them and pull out the ones we think are hopelessly flawed or not useful for the analysis. Before you do this, you need to have that upfront – the identification and quality stuff.

(48) **Question/comment:** Particularly for the examples. You're trying to test the reliability of the associations. I do think that there is a need for top down and bottom up when you are doing an assessment. I had a question about the biomonitoring. AOPs are descriptions of disease process. Biomarkers of disease are often late key events in AOPs. Those are the kinds of markers we would be looking for in epidemiological studies but that we're not necessarily seeing. To what extent currently are recognized biomarkers being measured in epidemiological studies? One other aspect, these patterns of empirical support are relevant not only to looking at epidemiological data but also toxicological data (perhaps in slightly modified format). We need to be considering patterns across studies; this is as or arguably, more important than the quality of individual studies, particularly those that will have no influence on outcome, due to poor quality or other. Poor quality studies are often more efficiently identified through initial consideration of patterns of results in the database, overall.

(49) **Question/comment:** I can think of dozens of examples that are controversial assessments going on right now that you can bring this method forward and help shed some clarity on the situation. I know one chemical in particular, styrene, where I know they have done a lot of work with this and there's a lot of controversy on it and a lot of epidemiology data and toxicology data and we do have adverse outcome pathway and biomarkers. That's just one example out of probably a dozen examples.

Response by Dr. Lange*:* I will follow-up with you on that. That would be great.

(50) **Question/comment:** The papers by Dr. Schulte I was referring to, you would be surprised to see the biomarker schemes bear a remarkable resemblance to MOA schemes as well as the AOP. It iterates in from external exposure to internal dose and then considers the mechanism and what is the actual dose metric you should be looking at. The "top down" adverse outcome referred to earlier is to the right of the AOP sequence of key events and the precursor key events are to the left towards exposure markers. There is a lot of math in terms of how you then verify a marker for those three bins: effects, exposure, or susceptibility. So, it can be one, two, or all three depending on the confidence that you understand that pathogenesis.



(51) **Question/comment:** That is kind of what we need. And the point to biomonitoring of exposure is the glass half-full or half-empty. It comes with perils. If you are doing it cross-sectionally or prospectively if you are doing repeated, there is a lot that as epidemiologists we hear biomonitoring and we say let's do that, it's great and then we get criticized because we didn't do it the right way. I think that is certainly something for discussion about not to waste the investment so that the results can be used as intended.

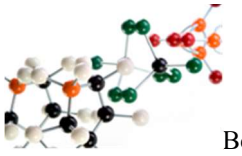
(52) **Question/comment:** It comes back to the biological plausibility and understanding the mechanism and not just looking at outcomes.

(53) **Question/comment:** And the dialogue would be great. Just an additional thought that you might consider on what we've been discussing around the context and MOA. There are a number of good examples and you touched on the one for your negative control in the pharmaceutical world. It's a classic case where we understand the mode of action in animals and how it happens in humans. We really understand the dosimetry precisely because it was done through Mayo Clinic studies and they know the dose of every single one of those individuals. Exactly the dose they received and no guessing. So, if you really wanted to look at some that are more precisely defined, I think you would find that there are multiple examples, positive and negative in the pharmaceutical world and I'm wondering whether that might help move your case forward. Because in the pharmaceutical world it is very common that they have the animal data on the mode of action, dosimetry, etc., because that is what the whole premise of the structure of the pharmaceutical world is built on. It is just a thought for whether that is another epidemiology dataset that could be further explore.

(54) **Question/comment:** And so, the title of your presentation was hypothesis-based. I didn't see a formal hypothesis testing framework and you didn't use the word Bayesian even once. Despite the fact that you are simulating a forward interference and hoping to use that to do the backward inference in evidence of patterns. You are simulating and the presence of a hypothesis we should expect to see these types of patterns? If you don't see them or you do see them, you're going to use that in your mental model. Is it entirely heading to the mental model process or is it going to be embedded in some other mathematical construct at some point? You have done the effort to do the simulation, so it implies it. You are not closing the loop, which is to say if I see that pattern what am I going to think differently?

***Response by Dr. Lange**:* At the moment, it is sitting in a mental model. Partially, I don't have the expertise to put it into a Bayesian framework without additional help and input on that. It would be nice to be able to incorporate it into a more formal mathematical method if we can and if it is conducive to that kind of an option.

(55) **Question/comment:** I guess my concern, and it is a concern on your behalf, I am concerned that all of the advice that you are getting is all very good. I don't disagree with any of it, but it is just going into a mental soup and you are creating a very complex mental recipe. The biological plausibility comes in. You cannot argue with the qualitative aspect of any of these arguments,



but what is the way that all of the evidence gets combined. All these inference arguments fall at the end to the fallacy of human reasoning. So, that is a concern on your behalf.

***Response by Dr. Lange**:* Yes, we need to be objective and need a clear path.

(56) **Question/comment:** But you are going that way. You are simulating the hypothesis evidence. But the Bayesian, it earns its way to if it's ever going to and going the other way around from evidence, observations of patterns, and back to the causality determination. That is the nature of my comments before the break. Yes, there is a pattern there, but if you saw it what would you do differently with that knowledge?

***Response by Dr. Lange**:* A good point and certainly something to think about.

(57) **Question/comment:** This is just a question that I should've asked earlier. I am assuming you are using this. One of the discussions we've had is about integrating data across streams of evidence. I am assuming this is integrative.

***Response by Dr. Lange**:* That is the intent.

(58) **Question/comment:** That is really critical. You are one step ahead as opposed to looking at individual studies. You really need to be integrating across studies. I think that biological plausibility can also be integrated.

***Response by Dr. Lange**:* Right. I very much appreciate the points but having seen this even with the control studies, now that we have this, then what? How is this helping us to make the final decision one way or another and to be able to function allies or operationalize that would be a huge boon for us, I think.

(59) **Question/comment:** I am no expert on Bayesian, but we have codified to what extent what we are doing. I do not know how you turn that into a high, moderate or low determination. How do you turn them into something that informs everybody else kind of thing?

(60) **Observer**: Going back to the causation discussion. I think it is much worse than we think. I believe this stuff and I'm not an epidemiologist. What to measure? We did a study using chromium and there were 10 different chromates. Each one was a different particle side. All we had was historical data for chromate ion. There is a huge misclassification right there. And another issue is if you have exposure data for chromate ion, what is the exposure metric that you use? Is it the default? Strong assumptions are built-in, for example, a reversibility. No dose rate effect. Short exposures at high levels are the same as long exposures at low levels if they have the same exposure and that is an assumption. There's plenty of evidence of attenuation. That is a problem. And then there is selection issues. They are different from the ones that leave. Sometimes they are different because smokers might leave faster if you are studying lung cancer. hat is a huge problem. There are all of these problems. This is starting to sound like macroeconomist s predicting interest rates. The Federal Reserve Bank makes decisions. They don't profusely apologize for them. They acknowledge they are uncertain, and they move on. I

think epidemiologists need some assertiveness training. We should tell the truth and not worry about it especially for risk assessment. We are making estimates at levels that are way below current standards often. We could be off by a factor of 10 and the conclusion would be the same almost.

(61) **Question/comment:** Thank you very much TCEQ for your case study.

Case study 4: Understanding Weight-of-Evidence of Ototoxicity from Co-Exposures to Noise and Chemicals in the Workplace

This case study was presented by Dr. Neeraja Erraguntla, American Chemical Council (ACC). The presentation is available on the ARA website. The purpose of this case study is to understand the Weight-of-Evidence of Ototoxicity from Co-Exposures to Noise and Chemicals in the Workplace.

Discussion and Comments from Panelists and Audience

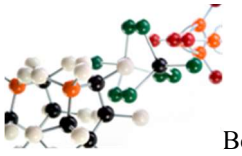
(1) **Question/comment:** Clarifying question. I was involved in some of the discussions with EPA when it came up with this really important thing to do and combined exposures, and there are no examples or case studies provided with any guidance.

Response by Dr. Erraguntla and her team members*:* There is information on lead and other things were discussed under combined exposures, but not on solvents. That's why we are calling it preliminary assessment. We only conducted an initial preliminary review of the weight of evidence and we will have to determine if we need to conduct a more thorough assessment. For that we will have to look into other resources

(2) **Question/comment:** Perhaps because Bette was involved, there was a direct mapping of the cumulative risk assessment guidance with AOP. There are some published examples of key events. Most recently, actually there is one linking aggregate exposure pathway, key exposures with adverse outcome pathways. Those explicit key events in both frameworks to do a cumulative risk assessment. So, I think you can use that same type of approach. It doesn't really matter. Both of those are stressors in the cumulative risk assessment parlance. Chemical and noise are both stressors in the cumulative risk assessment jargon. So, I think what you need to do is look a little bit more mechanistically at the damage. I saw you discriminate different types of ototoxicity, but it was always described as damage. You need to expand what the damages are because there might be concordance in terms of the spatial termination. A different type of ototoxicity might actually be the same type of damage in each location if you consider it mechanistically. So, I would encourage you to look at key events of what the solvents might be doing. But there are examples of cumulative risk assessment and consider them both as stressors and try to map them together. Have you looked at the literature from the Air Force on JP 8 and noise?

Response by Dr. Erraguntla and her team members*:* I did mention it. We have one slide that included that, but there is more information. We want to call this a preliminary assessment because there is a lot more to cover on this.

(3) **Question/comment:** The other advantage of utilizing the AOP is to look for other things that cause damage. I think the JP 8 story is pretty compelling that there is an interaction between noise and the components of jet fuel.



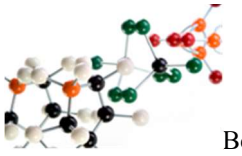
***Response by Dr. Erraguntla and her team members**:* For purposes of the case study and especially since we have limited time with you today, we were thinking that is a really complex mixture because in the component of the jet fuel, whether or not there are multiple components, we are trying to review single chemicals since the ACGIH published the Toluene: TLV® Chemical Substances Draft Documentation, Notice of Intended Change (CAS # 108-88-only for one chemical, i.e., Toluene in our case). Initially, we wanted to also review xylene isomers but since ACGIH's notice of intent we started reviewing the weight of evidence for Toluene.

(4) **Question/comment:** I think I was looking for that kind of model, because I can't imagine that EPA hasn't provided some case studies for cumulative approaches involving different types of stressors. I can't fathom a way to go about this without mapping it first and seeing where there's convergence of pathogenesis pathways. Otherwise, I really don't know how to move forward. Based on the interface of the AOP and systematic review communities, I am aware that the epidemiological community has been developing something they call logic models. These are essentially AOP type mapping but from the epidemiological perspective, with different terminology. They're addressing more observational empirical associations often of late stage disease in human studies whereas AOPs define more mechanistically based early key events. A kind of combined mapping across the biological/mechanistic and empirical/epidemiological domain might be helpful, as a starting point. I have a feeling you're going to have to rely a lot on epidemiological data here because it is very difficult to understand how, with two completely different types of stressors, we can incorporate or map animal studies exactly, unless they've been specifically designed to consider impact of these different types of stressors..

***Response by Dr. Erraguntla and her team members**:* We agree. If you look at some of the literature on the noise, we did want to consider the two stressors to cause the same effect. For this preliminary research, we did not look into AOP literature. We considered looking at it from the decision tree kind of model or review it as a hazard index and/or try something else.

(5) **Question/comment:** Just to clarify, for example, say there is a cytotoxicity. The degree of damage from one stressor or the other might not be the same, but that is the point of a mapping. You have got to know cytotoxicity, and the solvent might contribute some amount, and the noise might contribute a different amount. You can also create a network. So, the noise might be interacting with a solvent AOP at specific nodes. There is transition space between each of those key events. The noise might be modulating that. That is what we mean by mapping it out. It is not that you would need the same degree of the same key events. I just wanted to clarify that.

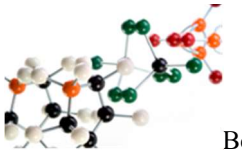
***Response by Dr. Erraguntla and her team members**:* Perfect. And we should say the NIOSH folks have been working on this issue a lot longer than we were putting this case study together. There are several representations of notation things, you can go chemical by chemical and see which ones meet with each mode of action or work from the mode of action and see what chemicals in the universe do that. So, you can address it from both ways. Even more complicating factor that I don't think we talked about today is noise is one of those stressors that is like a bona fide stressor. Like, if you make me put on earmuffs because I am exposed to noise, my stress response can happen regardless if I am actually experiencing noise or in chemical



exposure. My body and psychological reaction having to put on hearing protection can trigger that cascade. So, it is like one more facet of things that are difficult to put together.

(6) **Question/comment:** My thesis work was centered around ototoxicity, both from the epidemiological standpoint as well as developing an animal model from a toxicological standpoint. Noise is a systemic thing you need a whole animal to have and look at that response. And so, I think the animal data is strong and the epidemiological data on ototoxicity is very weak because you need such a panel of experts to pull together really strong epidemiological studies where you are measuring noise in the quality way such as with the information we now have on kurtosis, I think, is going to bring this issue much more to the forefront. We can measure it with much more quality. Then looking at interactions is so challenging. I wanted to summarize. Overall, this issue really brings noise research to the forefront because it is so critical if we are looking at epidemiological studies. I have a review article looking at different toxicological and epidemiological ototoxicity studies and different weaknesses that they have. You should check that out and NIOSH is definitely interested in collaborating. And we can definitely have further conversations on this topic.

***Response by Dr. Erraguntla and her team members*:** ACGIH has offered continuation education courses to talk about noise and chemical exposures. I am not sure if the target audience is e Industrial Hygiene professional or if they are including toxicologists and bringing them on board and talking. One thing I wanted to point out, as you go through it, is that toxicology is complex and depends if the noise is continuous or includes impulses. In one study that if the noise is continuous, it causes a positive effect. Where if it is impulsive, it causes a potentiate effect. So, interactions are complex. So, interactions are getting complicated. The doctor concludes that in her paper and how to explain that. What is the reason for that? So, these are a few other questions that we try to highlight. The different kinds of noises are important in studies when people are doing them, and we talked of the common nomenclature. It is not just to call it noise but you talk about the nature of the noise because that makes a big difference as a worker is exposed to the weighted averages. The length of that in the short term, how much is included? What kind of risk communication to make to the worker population and others? The other thing is that this is a concept, the startling noise can increase stress response. When you do studies and try to interpret that, how do you delineate those effects from what they are? That was one approach. So, thanks for your comments on that. The other approach is what we are talking about, the NIOSH approach in which they had looked into this. They call it an exploratory study. They have this kurtosis metric, a static metric that they take into consideration the temporal structure of the noise, and they looked at a study of 20 workers, workers who were exposed to solvents in China. I will let Krystin explain this if you want. I had to read more about it. This helped me. Work environments are not typically the distribution, it is a more complex distribution. So that is different. It is a statistical metric that is sensitive to the temporal characteristics of the noise. So, this paper I think is a way of looking at it as a preliminary way. Another approach, you look at the EPA cumulative risk approach. However, more guidance is needed to understand the complex interaction of a chemical stressor and a physical stressor such as noise.



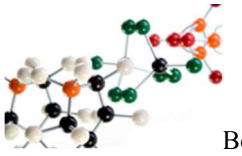
(7) **Question/comment:** I don't know what I can add to that. A pretty good description. We are looking at basically noise. It can be general, or it can be impact noise. When you look at construction work, people in a shooting range, or military exposure, there is a tendency for impact noise. In those cases, the exposure equipment is critical, and a lot of it is not measuring the impulse noise. If you don't measure the impulse noise, you are not getting the full picture of your exposure. That is what we are looking at here. Basically, I have an apartment on the third floor and my neighbors below me complain when my children are stomping their feet. That is the impact noise. But if I'm playing my stereo, it is just general noise. That was one of the things when we quickly looked through the literature. The impulse noise is difficult to capture in the cumulative risk assessment. It is a different type. It is not only a general organism stressor but a different nature of stimulus in the auditory and nervous systems. This metric is not consistent among the studies to do a meta-analysis or something more rigorous.

(8) **Question/comment:** What you're describing is, if you will, a dose metric. So, the kurtosis metric looks fine to me. We can try to map this, but we have a lot of different information. Has anybody conducted a specifically designed study to consider the impact of co-exposure? I get the impression, though I'm not familiar with the literature at all, that we're dealing, perhaps, with a complex issue. So, mapping of the evidence in the context of key events leading to combined effect, should as a minimum, identify critical data gaps. It also serves as an integrating construct to engage multidisciplinary experts in assessing available data and potentially, considering the path forward.

(9) **Question/comment:** I am struggling here as to a very practical approach to things. I appreciate the fact that one practical way of doing mixtures is the hazard index. Very simple, and we can do much better research. We can do mapping, pathways and get a much better way to go forward rather than just to a hazard index. But if you're talking about impact noise, it came to me to ask the question, do we do that for the short term exposure limits, do we do some sort of hazard index when we have multiple exposure limits, or do we always focus on a single chemical STEL and try to avoid that because ACGIH, NIOSH, and EPA all have methods for doing mixtures risk assessments basic to hazard index? Or if you have better information you can do it? But I've never seen that for short term exposure limits, which would be impact. Why would it be different? I don't know. You take the short-term exposure limits for two chemicals impacting the same organ system, and you don't want a hazard index more than one. Right? Is that how to practically do this? Is it so simple?

Response by Dr. Erraguntla and her team members: These are some graphs from the paper. I think we went over this that also pointed out the gender differences. We often hear the frequency is different between the two genders. Between different species too. One other thing that what kind of nomenclature we're talking about or consistency; even the regulatory noise levels are different in different areas and different zones and countries. So, a lot of things to take into consideration.

(10) **Question/comment:** Let me make sure I understand. When you say kurtosis, it is a spike of the noise signal. So, if it is a very general noise, it is going to be low over time. You're just talking



about impact noise. That is what you mean by kurtosis? Impact noise? What is a short-term exposure limit for the peaking of the noise? 88 decibels, a logarithmic scale. What is it?

Response by Dr. Erraguntla and her team members*:* The short-term limit is 85 decibels (dB) in the US. However, it depends on what country you're in. Talking of the US, the OSHA Standard 1910.95(b)(2) number is 140 dB at intervals of 1 second or less, ambulances 100 dB for 15 minutes.

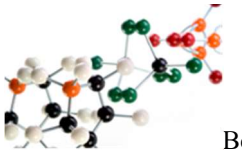
(11) **Question/comment:** I guess I'm going back to what the public is being told on their Apple watch. 100 decibels even if minutes a day can cause temporary hearing loss. The weekly limit at this level is around 20 minutes. Given all the literature, it depends on where you are. In the European Union, they have gone down to 80 as their noise decibel limit. Different durations for sure.

(12) **NIOSH Observer**: I joined late because I was just invited to the presentation. So, if I understand your question correctly, we have an impulsive limit for noise is 140 decibels peak sound pressure level and there are differences between 85, continuous noise, 8-hour time average and impulse noise. Just a follow-up on the kurtosis study, we are conducting a much larger scale study with our partners in China. We are collecting information on folks who are exposed to noise and solvents. Several thousand people. So, we are following up on that pilot study as well.

Response by Dr. Erraguntla and her team members*:* Thank you. That is helpful. Again, that is something we learnt from NIOSH, and I'm not going to go into the area, you are the experts, but that is why it was getting complex. We started this project and quickly the literature was getting complicated. The NIOSH reference was looking at the 3 dB exchange rate. And that is defined by OSHA default with 90 decibels for eight hours. The noise literature is pretty complicated, and the NIOSH experts evaluate each study. When we reviewed it, for example, we had to go and see what the noise was without blindly doing it. We needed to be very careful. So, we thought we would start tabulating all the information into the evidence table to be easy to look at what the evidence is saying before we can go further. In the context of impulse noise, short-term means really short-term. It is like 25 milliseconds. That is the duration of impulse. It is not integrated or for 15 minutes. Really short. They are big and rare, but not so rare they average out. That is why kurtosis comes up.

(13) **Question/comment:** That might be a great chart we can collaborate on to make sure the terms of reference basically as we advance forward because in a case study setting, trying to put up an audiogram of about 20 printing press operations and solvent exposure and then the second career was training as a police officer. So, to the extent that for any individual, there is research, one bullet single gunshot equals one week of hazardous occupational noise exposure. So, I just think anytime you go to a shooting range, are you shooting up your risk?

(14) **Question/comment:** Just a question. When you are talking about the differences in the frequencies, with therapy is there a different threshold potentially for different frequencies?



Response by Dr. Erraguntla and her team members: That is to come. That is one of the struggles we were having. But obviously we don't have time to review the literature, but we need to do that. Some of the literature in animals, different solvents, some are biphasic, some are the hearing loss depending on if the sound is in the mid-range, some of the high range. In humans it depends on gender. I am assuming in animals it probably does too. Which one is more sensitive? I don't know. Depends on what you value. And not to understate, the military has done a ton of work in the space through the VA. Hearing aids and overcoming tinnitus and advances beyond just a simple hearing threshold. There is a lot more tools we can probably use as we tried to bridge that gap. But even single solvents, single chemicals in an animal, it varies depending on the frequency, varies depending on the duration and repetitiveness.

(15) **Question/comment:** Trying to push this agenda a little bit on the earlier question because we're trying to make this a tractable problem. Have you tried to do a crude hazard index just to see the extent of the issue you are dealing with here?

Response by Dr. Erraguntla and her team members*:* We started off as a case study, but we quickly realized that we didn't have the time for this.

(16) **Question/comment:** Say, that you just take a measure for the two chemicals and the sound. Can you do a crude hazard index? I'm just trying to get a feel for it.

Response by Dr. Erraguntla and her team members*:* Definitely. es. that is something that our group might consider for some chemicals of interest and as part of a broader collaboration with multiple stakeholders. Jim and others have authored a nice paper with a big section on ototoxicity for styrene. Our group reviewed this paper.

(17) **Question/comment:** But you can do simple hazard index and add chemicals in to see the impact.

Response by Dr. Erraguntla and her team members*:* Yes.

(18) **Question/comment:** The exposure samples to styrene, particularly in the fiberglass manufacturing industries, are almost exclusively to styrene. One of those rare occupational scenarios where you just have primarily exposure to the one compound, styrene.

Response by Dr. Erraguntla and her team members: It was a timing issue. I actually reached out to the group to see if we could collaborate with them. They could not meet us, but perhaps in the future we could potentially collaborate.

(19) **Question/comment:** The other thing I should mention with styrene. The current OEL levels are pretty much set around ototoxicity issues because there had been relatively robust occupational ototoxicity studies conducted with styrene and West and East Germany about a decade or so ago where they latched onto a cohort of workers and were able to very precisely define where the specific hearing deficits were and at what frequency. And ended up ultimately with the OEL value being dropped to 20 from 50 at the time. Because of that recommendation, what the overlay of sound was on top of that I don't know. The 20 value, just measuring the outcome,

hearing loss, it is inclusive of sound plus styrene and ends up at 20 parts per million. The question you're asking if I hear you correctly, how much additional impact to add to that? I am not sure. I am not sure how profound it would actually ultimately be.

Response by Dr. Erraguntla and her team members: Back to your question, Bette. The Australian number, in Europe I think, it is 50 ppm. Here 20 ppm. One of the things we also are looking at is ototoxicity happening with the 20 ppm. But now we're bringing noise. That is something that our evidence tables, I think, will have. A little more to come. Definitely.

(20) **Question/comment:** I do not know if you have an answer to the to the question at the top of the screen. The CNE metric for kurtosis.

Response by Dr. Erraguntla and her team members: Yes.

(21) **Question/comment**: OK. I don't know if you have analyzed the Fuente's paper distributed with the case study. Is that contrary evidence to that? What company has tried to replicate it and found nothing?

Response by Dr. Erraguntla and her team members: It was an understanding that that was an explanatory study for a small sample size. That is what we heard. The larger sample size might get us more on this. Other than the paper, I did not find anything.

(22) **Question/comment:** It does have a little bit in the paper of mechanism as well. Is that one the people agree with?

Response by Dr. Erraguntla and her team members: We just started this discussion within our group and started reaching out to other groups to have this discussion. So, the timing of this workshop would definitely encourage more of these discussions.

(23) **Question/comment:** I can only say having looked at it as quickly – not even as long as you have –, that it just sounds like it is inherently reasonable. Right? The intensity is a culmination of dB, duration and how hectic the sound is ultimately. How much more than what people would normally be exposed to? It seems that inherently has good ingredients. In a little bit like what we were talking about with exposures, which combine low concentrations for a long time. It is an assumption one has to make, and it just seems to me to be in that same category. It maybe there for another decade.

Response by Dr. Erraguntla and her team members: I think that big study referred to will definitely add to the subject. There were a couple of questions we had in the paper that had not one solvent but four or five. And then they came up with a way of combining them. We will definitely have future talks with the NIOSH colleagues and see how to understand that and wait for more results. Thank you.

Response by Dr. Erraguntla and her team members: We have one last thing to share. We reviewed a journal article that used a biomarker to make correlations with hearing loss I found

this definition, of a biomarker and the purpose of including the definition was to put it into context that. I'm sure there are others in which a biomarker is a substance or structure that can be measured in the body and predict outcome of a disease, but to be clinically relevant, the biomarker should be specific for that certain disease, and correlate with the specific s activity.. People are using the data and making causal correlations without strong associations. You might have correlations without strong associations. In this particular case, the biomarkers used were oxidative stress. This is a busy slide, but this was a study I found. A number of papers have been published on oxidative stress and associated oxidative stress with cancer. But these are all the different target organs and different papers. A tremendous number of papers have been published on oxidative stress. You just Google and you can find them. That was one thing looking at the study quality. Have the authors done their homework in looking at it as biomarker of exposure or is it a biomarker of effect? We just wanted to have the discussions here. They were looking at metabolites, oxidative stress and just went and made correlations using oxidative stress and correlated with hearing loss, for many chemicals, parent compounds, and metabolites. The purpose of this including this is to show that oxidative stress is a non-specific biomarker for hearing loss. We wanted to bring it to your attention that people may continue to use it to make other making associations without strong evidence.

The nice thing about the paper is in the that they included audiometric hearing data from 2012 and 2011 NHANES data set. They had some information on the metabolites and also information through audiometric came up with this correlation. The other things, this paper has been published recently. They have reported that this particular auditory hearing has some limitations. It is nice that they explained what the limitations are. Whereas the other paper we were referring to did not do their due diligence of explaining what the limitations were. Some of the things we hear in the literature, lack of validation correlating the biomarker to disease. Disease in general. In this particular case, we are talking about hearing loss. The question to the group, to the panel was, what do you think of oxidative stress and the biomarker. And that can be used for auditory hearing loss. Basically, get your opinion as to using this biomarker for hearing loss.

(24) **Question/comment:** A comment on oxidative stress relative to hearing loss. Personally, I would put that as an extremely low probability that you would ever establish any meaningful correlation. The reason I say that is, think about it, what is the tissue mass of your hearing organs, relative to your whole body, and relative to oxidative stress that is also occurring throughout your whole body? I will guarantee you the variability in your oxidative stress just associated with your normal, whether I'm standing, sitting, decided to walk over lunch, will swap out any oxidative stress that could be contributed by probably just a few milligrams of tissue relative to the kilograms of other biological tissue that I carry around with me every day. I would be stunned if oxidative stress in terms of a biomarker in humans, and it would normally be measured by urinary products, could ever be correlated. It might be a measure of your stress level. If you are in a hearing range, and somebody was pounding on your -- the apartment above you, you are getting impact noise, that is stressful, but the question is, is that going to -- is the oxidative stress that might appear in your urine as a result of that? That may have nothing to do with the impact on your hearing organs whatsoever, but it does have to do with your physiological reaction to it.

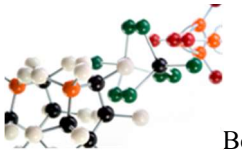
(25) **Question/comment:** I guess that would be another nuance of even the enhanced study. Trying to get every bit of relevance out of any study we run across. To the panel discussion and the other presenters discussion, earlier today about what is adverse, if we work through it for hearing loss or these markers in the context of hearing loss or in the context of solvents or toxicity or localized or whole body affects, the interesting thing is that you are still going to keep getting correlations. It was like a 3 or 4 decibel shift in regular audiology exams. It wasn't age related, and it was in people who reported no exposure to noise. So, it is putting together each of those pieces and then kind of presenting them as, is as part of the mode of action or not and taking that and saying if the auditory stress, oxidative stress are not part of the mode of action relevant to address this issue, then take it to the other extreme. What about contribution to different disease states? And so, I guess that is the interesting thing for me personally. My new study area is on dementia only because my parents have dementia now. So, I am trying to figure out how to best help them. Hearing loss, depending on what country you are in, South America and beyond, 5% to 7%, hearing loss and middle age is a risk factor for dementia. Is there something in common, either the stress pathways, the metabolic pathways, genetics, what is that little nugget that might be in common? And we end up saying actually hearing loss has nothing to do with it. It is all about the pathway.

(26) **Question/comment:** I'm wondering if we are not constraining the definition of a biomarker. In the case of hearing loss, that is a biomarker. Is it not? It is a measure of function that I can run a test on, a subjective test on to determine hearing loss. It is no different than if I am measuring a change in and oxidative stress biomarker in my urine. In the case of hearing loss, the question would be, why do we even have to mess around looking for tissue or urinary biomarkers when you've got the biomarker at hand, which is well-defined and very sensitive in terms of measuring hearing loss, even across very specific frequencies? It is got all the dimensions of being a biomarker of hearing loss. It is capable of studying interactions, weather sound or chemical or stress or you name it, the biomarker is there. It is the testing; the test is not complicated to do necessarily. It takes only a few minutes to get it done in the audiometry laboratory. So, should we not just be defining the biomarker for hearing loss as hearing loss itself?

Response by Dr. Erraguntla and her team members: I like that. It is a nice way of looking at it because people have not been considering hearing loss as they start to figure out the sleep disturbance and that causing other cardiovascular and other diseases. We just did not call it a biomarker. So, you get the credit.

Response by Dr. Erraguntla and her team members: I agree. I think one analogy one might consider is pulmonary function; it is the same.

(27) **Question/comment:** It is, except for that it is also an end disease. It is treated by physicians. So, you are prescribing a medical device to treat the lack of hearing and overcome that inability to function in daily life. So, to me hearing loss is a disease. I think you can use pulmonary function in various ways to diagnose disease.



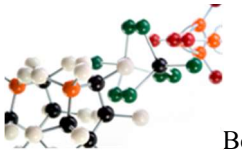
(28) **Question/comment:** I wanted to introduce the complexity to the simple audiometry because there is a lot of discussion on that right now. In the forefront of hearing research, it is actually reporting there are a lot of people that have what they call hidden hearing loss where they can hear well in the audiometry booth, which is soundproof, however, when they go to a café or they are going somewhere with a lot of background noise, there is a lot of neural functions going on to dampen this background sound so you can have a conversation with the person next to you or in front of you. That is something that is really difficult to hit in the audiometry. There is also distortion product otoacoustic emission (DPOAE) testing for the outer hair cell function. Then if you go into further testing, you can look at the brain waves as you see in audiometric response (audiometric brainstem response or ABR), and you can analyze the amplitude of those and the distances between these waves. So, one of the problems in looking at the epidemiological research as well as the animal research is that the outcomes are in different metrics in juxtaposition to the exposures very difficult and complex.

(29) **Question/comment:** I am glad you made those observations because I would submit what you just described are exactly the type of where you are likely to be more productive in terms of identifying biomarkers for evaluating the impacts of solvents and sound on hearing loss What I would be concerned about is if you go down expecting you're going to identify a biological biomarker in terms of, let's say, I can detect something in my blood or urine, the conventional biomarker used. It is likely not to be very productive. But if you use the biomarkers you just described whether it is audiometry or refined audiometry read the background noise on top of the standard, I am sure that can be done. Or, whether you can use whatever technology you have, to me, that is much more likely to be productive in terms of leading to improvements and understanding of how these interactions between sound and solvents etc., where they begin to interact. I would just be concerned if we have the expectation, you're going to find a magical biological biomarker in our conventional definition of it. It is highly unlikely to be successful.

(30) **Question/comment:** And the question of the weight of evidence. You have evidence from the different streams and how you integrate the data. So, I think in the interest of time, we just put this question out there.

(31) **Question/comment:** I am trying to understand what exactly we already know. Do we know that in the combination of solvents and noise exposure you would have lower acceptable exposure levels? Is that correct? Do we know that for sure? We only have the question of quantitatively getting the level when you have those two stress factors that you need to identify the safe levels.

Response by Dr. Erraguntla and her team members*:* I guess what I was going to say is, it depends on the context. Comparing to OSHA in 1973 to 100 dB, it probably needs to be evaluated to see if it needs to be lower than that based on the animal data and human data. Based on the JP8 jet fuel studies. There are a lot of animal studies on this topic and these have been conducted at relatively higher dose. I don't know that we definitively answer that for jet fuel and noise in combination. I think that is something important the DOD is working on. For every individual chemical, the answer might be different. They are not all the same. The threshold at which ototoxicity evidence starts, it is not an automatic across-the-board concern. That is exactly the point at which we are. We think the evidence right now says we don't need to go



lower based on the evidence and also NIOSH has a large ongoing study on this topic. However, that is something to be confirmed and laid out after systematic review and quality review of the literature.

***Response by Dr. Erraguntla and her team members**:* ACGIH has a notation for toluene. However, there is no guidance and the question is will the notation by itself change the behavior in the workplace.

(32) **Question/comment:** What I heard is that in the animal studies, you don't have a lot of good exposure information of the confounding factors to clearly determine the ototoxicity. But in the human studies, like occupational exposure, you do see that there is likely that the noise in combination with exposure, you would have potential lower safe levels. I was trying to understand, based on the understanding that what you see is likely protective already. Is that right?

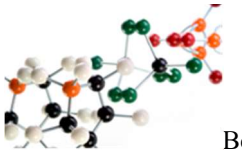
(33) **Question/comment:** I would agree with that assessment at this point. The flipside is that the noise level could be reduced and the issues remain the same. There is sort of a mirror image associated with this kind of study. For example, in Australia and Europe, OEL was 50 ppm. And the noise level was lower. Here, the tolerant OEL is lower and the noise level is 85 dB. So, it is almost like a little bit change on both. There are several reference levels: Occupational Safety and Health Administration (OSHA) TLV 200 ppm, NIOSH PEL 100 ppm, the ACGIH TLV 20 ppm. And having the OTO notation helps bring this awareness. It could be different for different solvents or chemicals too.

(34) **Question/comment:** What did you find in the literature, because you said OEL might be protected? What about co-exposures to multiple toxins plus the noise? Is there anything in the literature to give guidance on that?

***Response by Dr. Erraguntla**:* The kurtosis paper, they looked at four solvents.

(35) **Question/comment:** It is not my expertise, but I may be able to answer that. In France, they developed a software system, a program. They can input multiple chemicals and substances and, we are trying to see if we can mix and add noise to it. I don't know if I answer the question. For folks, for the question, the paper I think has a question to describe how to get to that. This is just a clarification. It is calculating a hazard index I presume. Just doing the math. A spreadsheet. Did you look into the literature that has the multiples solvents?

***Response by Dr. Erraguntla and her team members**:* Like we said, when we start looking at ototoxicity literature, we come across many of the chemicals and some papers that talked about other solvents. We did not look at all of them. Just reminding us it is there already. Some of the studies include multiple solvents. For us, maybe this is a preliminary thing. Like we said, when we start looking at ototoxicity literature, you come across many of the chemicals, but we try to restrict, but we have some papers that talk about other solvents. We did not look at all of them.



(36) **Question/comment:** This is not my area of expertise, but I will jump in. I really encourage this idea of collaboration. It is complex. That is really pretty obvious from this. But there are ways to go forward. We have authors in the audience here that can be pulled in to help a little bit. We had the publication that was referred to several times about the tiers. If you're going to go for a hazard index, you don't have to, but if that is your formulation, is the OEL low enough or high enough with noise exposure, started zero, they go to one and add complexity as you need and add complexity on the exposure side for either chemical. But the idea is really to foster and encourage collaboration because there are a lot of smart people here, and I think you can solve that problem if we work together.

Response by Dr. Erraguntla and her team members*:* That is how we all started. We started working on this case study during December 2019. This is all we can do at this stage with a month ago, a preliminary
assessment.

(37) **Question/comment:** I just want to comment back to the analyst, you can consider JP 8 and add both mixtures and look at them and unpack them if you want. But certainly, they are mixtures in of themselves and have demonstrated to have interaction with noise.

Response by Dr. Erraguntla and her team members: We also try to look at one of the things that came out, I think, in this workshop, a previous one, looking at a visual graphic to see where you start, where the OEL is and kind of build it out. So, we started doing that. We just ran out of time for this. But that is something you can unpack and see and visually put it out there. We try to do it with that one slide.

(38) **Question/comment:** I know Dr. Dave Matte just up the street from us at Wright-Patterson AFB has already done all that work. I was just advocating that JP 8 is a great consideration since they did parallel animal and human studies. They have amply demonstrated hearing loss and JP 8 interaction up the street.

Response by Dr. Erraguntla and her team members: One thing quickly that happened is we started doing the outreach to the other groups working on. Wright-Patterson AFB also has nice hearing loss equipment. You might want to make note of that as well. I guess that is among the collaborations where we are not represented today unless the Army has guidance out, they have had it out since 2003 doing hearing loss threshold shift audiometry also consider co-exposures to the fuels and that other stuff. There are tons of military researchers working on it, and like I said, the VA is basically on top of it for purposes of treating and intervening earlier. That is really the most beneficial result for all of us, to figure out and make sure we're getting the hazard identification right. Is it in both of them, one of them, none of them, or combination only? The other thing I think Krystin pointed out, the audiogram that NIOSH is looking at, and the newer ones coming. Also bringing the workers medical history and see what their background exposure is.

(39) **Question/comment:** Just another general question. In general physicals, especially for older people, who may have had an accrual of a variety of pharmaceutical interventions over the last

10 to 20 years. There are a lot of dermal screens now that are a major part of a regular annual physical. Is it a situation where hearing tests are part of a regular general physical and is there such a database? There is a U.S. Task Force proposing to add hearing loss as one of the – I don't know what they call it –, medical indicators for the Asian population. It is in the works. We just submitted it, not us, but as part of CDC some comments and reviews for that proposal. It is in the works.

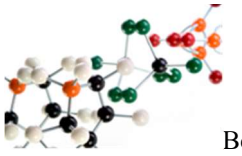
Response by Dr. Erraguntla and her team members*:* Thank you.

(40) **Question/comment:** That brings up the medical history thing. There is a lot more happening on the ototoxicity monitoring program. We've looked at them and need to get much more. Other data requirements based on what is available there for the JP studies, maybe controlling and accounting for the confounders. We definitely heard of oxidative stress. Not a specific biomarker. We can take that out. As a buildup in the motive action, what comes out, the next validated biomarkers. What else should we include? We talked about the possibility of the ototoxicity and neurotoxicity connection. Thinking about it in terms of using animal testing, how would you define, how would you look at new methodologies for that? What should be included in the chronic? These are questions to the committee as to what your minimum data requirements for what should be? Obviously, we would definitely like to have uncertainty and sensitivity analysis included.

(41) **Question/comment:** This is a question to Jim or any other panelist. Do we do testing or do we do auditory function? Do we even do auditory histology?

(42) **Question/comment:** Not the testing. It is not full range. That is given at a fixed frequency. I do not know how comprehensive that would be relative to the problem formulation here.

(43) **Question/comment:** This summarizes again the next steps forward for the next workshop. The next collaborative project between different groups. Thank you everybody for your valuable comments.

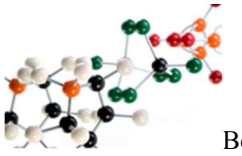


Case study 5: Risk/Benefit Methods for Carcinogenicity / Sterilization with Ethylene Oxide as an Example

This case study was presented by Dr. Lucy Fraiser, Lucy Fraiser Toxicology Consulting LLC. The presentation is available on the *ARA* website. The purpose of this case study is to compare theoretical cancer risk estimates from exposure to EtO concentrations in ambient air near medical equipment sterilization plants and the countervailing increase in the risk of healthcare associated infections (HAIs) that are expected if EtO becomes unavailable for sterilization of multiple instruments simultaneously in procedure/surgical kits.

Discussion and Comments from Panelists and Audience

(1) **Question/comment:** That was a really good presentation. Thank you for that. Ethylene oxide, the issues you brought up in the slide were strongly debated in 1983 and 1984 at EPA. During that period of time there were three agency statisticians who worked vigorously on this issue of how you do the extrapolation, what kind of mathematics occurred. The consequence of that decision alone is really quite spectacular. There was an agency policy decision to use the supralinear even back then, which caused a huge amount of scientific discomfort., but it solved the policy problem so to speak. It might be helpful if you resurrect some of that, because I don't know and have not tracked the math conversation. However, if it still stands that he way you get to the answer goes through a policy decision rather than a mathematics – as legitimate as it may be for a regulatory agency, when you do a risk-risk evaluation, it might be helpful to show the impact of a policy decision on the answer for real risk when you are doing this kind of thing. I am not suggesting that EPA should not do it that way, don't get me wrong. When you look at actual anticipated real-life risk, particularly in these kinds of settings, one of the impacts on the answer is a decision how to draw the line basically. It gets down to that, and it was a significant difference. You might want to resurrect, if you can, some of the conversations about the alternative ways to come up with the cancer risk number. I am sure there is no way in this stage of the game you will solve what the true question is, which has been debated for lots of chemicals for a long time. Nevertheless, it's relevant for this conversation. Another issue that is related is the actual risk to either ethylene oxide treatment of the medical devices extends beyond the operating room. These people don't just get up off the table and go home with their injuries fixed. There is going to be a barrage of additional instrumentation which will be applied to these people, whether it is a drip line or other insertions. All of those are ethylene oxide treated as well. There may well be statistics on that kind of thing. If you think about every time your blood is taken or every time a nurse hooks up a saline drip and opens up the packages, those are ethylene oxide treated as well. I don't know what impact, but it seems reasonable that a postsurgical patient would be expected to see these products multiple times. To other points, if you substituted another sterilant, what is the chance that the risk would be zero? Either to the patient -- if you come up with another chemical, you are looking at another risk assessment. If you have a compound that poses no biological risk, it is likely not effective on it as a sterilant. So, because of the nature of the chemical, part of the risk-risk decision is an unknown. The idea that could be zero is not on the table. It also poses a systemic risk to the hospital worker if they have to now intervene, where they do not have to. If the nurses have to remove something from



a constant chamber where you have an atmosphere with surgical clamps or Band-Aid or gauze, you're introducing a new risk factor to a new group of people. Back then, that was one of the conversations that we had. Who is the person who is being exposed to some of the alternatives that were in play at that point? These are just ideas that had been discussed in the past and I hope you can get a hold of some of that literature. Yes, I mean the field is wide open.

***Response by Dr. Fraiser**:* In fact, you can find some older studies where I guess hospitals did have ethylene oxide sterilization chambers inside the hospitals. The cohort are nurses.

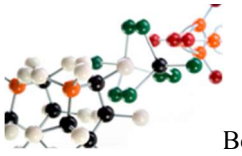
(2) **Question/comment:** Thank you for your presentation. I was part of the EPA advisory board in 2015 and attest to the fact it was controversial at the time. It was not a unanimous decision to go forward, although it was voted forward. I can also attest that the EPA did not address all the comments in their final. Having said that, I am also now working with industry in Georgia to do monitoring which leads to my comment. If you look at the cancer risk, you are using an upper bound and usually it's the lifetime average daily dose assessment, which is about four-fold lower than where you are. That's maybe something you should put in as an addition when you do your theoretical cancer risks. The fact you used TCEQ is a good idea because EPA did not have full knowledge of endogenous levels of ethylene oxide. Our bodies make it every day. It was not reported in 2015. TCEQ is gone further along that line. If you say that the safe dose or up-to-date risk specific dose is somewhere between those two, that is probably a reality. Thank you for the risk-risk rate up. I enjoyed it and I have specific comments I will get to you which are minor.

***Response by Dr. Fraiser**:* Thank you.

(3) **Question/comment:** So, I have full disclosure that I am an advocate of data. I don't know who the champion is for collecting these data. One question or suggestion is if the data around air monitoring and sterilization plans is inadequate, are they motivated to collect their own data around their own facilities to improve the status quo? It's important and reassuring if they do. I heard there is a huge shortage of medical supplies in Cincinnati and it's a problem. When you are talking today, I wonder if there is an opportunity and who is a stakeholder and is motivated? Do you have hospitals right now that you can collect real-time data? Are the infection rates changing based on shortages of availability of these kits?

I am thinking the study hasn't been done yet. The question is, are they at a point where the data needs to be collected?

***Response by Dr. Fraiser**:* I looked for studies. There are a lot of studies that looked at overall infection prevention program impact on infections. There are many studies. Every year of National Healthcare Safety Network data are updated, and it seems the rates are going down quite a bit. A lot of the rates published in the literature that focus on older data are a lot higher than what you find in most current and updated data. For one thing, insurance; they are starting to not get paid if an operation or procedure causes an infection in the patient. Unfortunately, I did not find any study that focuses specifically on the use of these surgical/procedure kits. My client that I am working for was the one who suggested that you need to look and see if you can find data on decrease in infection rate since we started using the prepackaged kits. Studies attest



to the fact that they decreased time in the operating room (OR) and mistakes. Both of those can indirectly reduce infection risk. But I did not find any studies that focused specifically on the kits which is why had to come up with the calculation which makes it a lot harder. I wish the data were available.

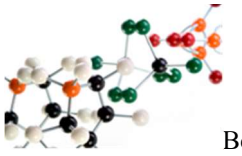
(4) **Question/comment:** I'm wondering. They may not be in the published literature. It may be the time to search out these quality data that hospitals have to collect and don't have to promulgate. Is there data, probably not that you collect, but a motivation for someone to collect better data?

(5) **Question/comment:** I was a co-author on the 2009 study that looked at the heritage Union Carbide data. I saw Gary Marsh did a meta-analysis in 2019. It doesn't look like there has been much done since the last paper I saw in 2011. Do you have a view or position on the data to date? I don't know if you have a care at all but wondering what the view is on ethylene oxide and the cancer?

(6) **Question/comment:** We have ethylene oxide cohort which has been studied a lot. We have been looking at it recently for issues relating to things like survivor bias. We worry about extrapolating down to a one in 1000 risk and one in 100,000. What does that mean? On the other hand, linear exposure response is a good default assumption. I'm not familiar with the EPA treatment with that. I don't think we have anything to add to this. One thing we like to talk about is engineering controls. My question is how is this stuff getting out to the environment? Do they vent it out the top of the stack? Do they not scrub it or catalytically treated or anything?

Response by Dr. Fraiser: They are treating it and in fact one of the things going on now is, there are a lot of things that happened in Illinois where the Illinois EPA issued a Seal order and then the plant decided to close its doors. After that, we have seen plants in Georgia that voluntarily closed down so they could add on more controls. It is getting a lot of attention, and the communities around these plants have become alarmed because they don't put these plants out in the middle of nowhere. They are across the street from police departments and other city buildings. The communities are really up in arms and the plants are adding controls. They get pretty incredible ethylene oxide emission reductions, before it gets out the stack, like 99.9%. I am probably talking outside of my area of expertise, but one issue is, you have fugitive issues. I think that is where we see the variability.

(7) **Question/comment:** Thank you for that intriguing presentation. It's always important to consider the countervailing risk as you have done. One question I had, and you touched on it for sure and you are probably thinking this way already, but as you talked about the potential disadvantages or issues associated with your analysis, a significant portion of your analysis rests on the SSIs (surgical site infections). If I understand that correctly, maybe I am just defining it too narrowly, but when I think of SSIs, I think of surgical sites. In those scenarios you had the assumption there was 85% cross patient contamination, but my guess would be in a surgical site scenario, the cross-patient contamination is essentially zero because what is in there stays there. One patient gets operated on and all the equipment goes in the trash afterwards. Where the cross-patient contamination comes in and we have seen it is on the floor. Nurses and doctors go from patient to patient and from room to room and are too lazy to change the gloves and not

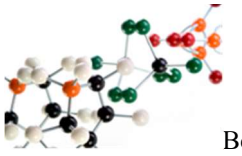


disciplined enough. I'm just wondering if you thought about ways to approach sensitivity analysis in the methods you have done. If your method currently now is heavily weighted to the SSIs that are restricted to surgical site things, then your assumption of 85% contamination is too high. Then you could evaluate that in terms of the impact of your calculation. By the way, I should mention as I am commenting on this, my employer and myself have been actively involved in the ethylene oxide issue as a consulting company so I should make that transparent. That's the comment I had. With respect to the comment with respect to occupational exposures and ultimately plant releases, interestingly from what I understand about the sterilization plant operations, for a long time they have been operating with really restrictive PPE protection for the individuals in the plant. So occupationally, these individuals operate essentially with 100% protection. This is full masks with air-breathing scenarios. The opportunity for them to experience exposure is pretty low. The plants have been robustly treating their exhaust emissions and continue to refine the technologies. Most plants around the U.S. are undergoing additional refit's that will further reduce emissions from those plant operations. One of those things that has been treated as a result of recent EPA efforts, you mentioned motivation to do testing on the part of the manufacturers. All of that is happening. State governments are doing it as well. One of the intriguing things that has come out of those observations which was not generally appreciated before this episode came to the surface in the last year and a half is, there has been a substantial and variable amount of ambient air in ethylene oxide in the air today. The air we breathe today is in the range of 100 parts per trillion of ethylene oxide which is a 10 to the minus 3. The sources of that are not understood and could be a multiplicity of sources. And your own body, as mentioned earlier, is estimated that you produce the equivalent of a two part per billion exposure to ethylene oxide which is 2000 parts per trillion of ethylene oxide per day as a result of your oxidative metabolism resulting in metabolizing ethylene oxide which is how that happens. My primary point is as you think about diving into your method to say as you would do with PBPK modeling, for example, try to think what are the target sites for appropriate sensitivity analysis to say what would be the variables that would impact my risk decision one way or the other with respect to this infection potential.

***Response by Dr. Fraiser**:* All right. Thank you. That's a very good comment. To me it seems 85% is high. A lot of the statistics that are reported in the literature about infection rates, they are really high. It makes you a little paranoid to go have a procedure.

(8) **Question/comment:** I really enjoyed the case study and great presentation. I am just curious about the statement about there being no medical device that has ever failed to be colonized with coagulase negative Staphylococcus. I was thinking of all my friends who have had knee replacement and hip replacement, and wondering about microbes present on the medical device surfaces after ethylene oxide treatment? What is the reference?

***Response by Dr. Fraiser**:* The devices themselves don't have microbes on them after ethylene oxide treatment, but the two studies that looked at device contamination look at the packets themselves and at the packet opening process. You might have devices that sit around and have particles drop on them. Or the package gets contaminated because someone's gloves are contaminated.



(9) **Question/comment:** Okay. But now looking more broadly at the surgeries that people are routinely getting and implants of medical devices, I would think there are far more than just a central line contaminations.

Response by Dr. Fraiser*:* The case study also included surgical site infections. That would include the orthopedic surgeries. It is a big source of infection because one of the practices in orthopedic surgeries is that every screw they use would come individually. The studies that looked at those risks were associated with orthopedic screws used in those types of surgeries.

(10) **Question/comment:** I would think there would be some data on complications of those surgeries even though you don't get the reports when you are discharged from the hospital. Although I know a lot of people who have had no complications after a knee replacement, I know one person who had chronic infections that just could not be cleared. So, that is almost a lifetime disability associated with potentially errors in the sterilization procedures. That kind of error would be increasing if you didn't have access to ethylene oxide anymore. I would think you might be able to make a case for more than just acute event in the microbial side of your assessment.

Response from Dr. Fraiser*:* Thank you.

(11) **Question/comment:** Regarding the slide that showed that the overall equation a series of variables, can you describe that exactly what units that is in?

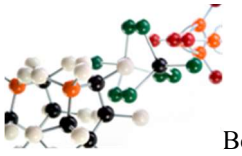
Response by Dr. Fraiser: Those are the data that are reported to the National Healthcare Safety Network on central line associated bloodstream infections and surgical site infections. Basically, they come up with that number by taking the observed number of CLABSI divided by the central line associated days. You often will see them reported as CLABSI's per 1000 central line days. I used just the rate divided by the number of central line days and that way you get a probability. For CLABSI, I believe, it was 7.9 per 10,000 central line days or something along those lines.

(12) **Question/comment:** So, it's not a number of infections. It is a rate?

Response from Dr. Fraiser*:* Yes.

(13) **Question/comment:** Okay. In each of the other ones is a probability on a per procedure basis, on any given procedure. What is the probability that procedure will result in infection?

Response by Fraiser*:* Not exactly. For example, the probability of infection caused by a specific microorganism was based on looking at what microorganisms are reported to cause a certain percentage of infections. That would be what percentage of CLABSI are caused by coagulase negative Staphylococcus. What percentage of CLABSI are caused by Staphylococcus aureus? So, it's not a procedure but a percentage overall. So, it's a fraction caused by each type of microorganism?



(14) **Question/comment:** And then the next one, the probability that the colonization progresses to infection is that on a per procedure basis?

***Response by Fraiser**:* It is not per procedure basis. In some cases, like with coagulase negative Staphylococcus, I was able to find data on colonization of central line insertion point and what percentage of those progressed to CLABSI. I was not able to find that in all cases. In some cases, that represents the fraction of commensal colonization. The natural colonization of a patient's membrane or skin and what fraction of those patients develop an infection. I tried to focus on the specific type of infection, but I did have to use commensal colonization data in a number of instances. And then the others are simply fractions or percentages.

(15) **Question/comment:** I guess what I'm trying to understand is like with the infection rate (IR) variable, you already have infections. Then you are multiplying by the probability of infection which seems like you are double counting the probability of infection because it's already in the baseline data.

***Response by Fraiser**:* But I am trying to estimate the increase in infection if ethylene oxide is banned, so we go back to a scenario where you are opening multiple packages of medical devices. This is the increase above what normally occurs.

(16) **Question/comment:** But isn't the baseline situation you are describing where the kits have multiple parts -- is that not part of the baseline data? The status quo, is it not for a lot of procedures to use the kits?

***Response by Dr. Fraiser**:* It would be the status quo for a number of procedures. What this is estimating is if the procedure kits go away which is part of the baseline number and we start having to open individual items, so this is the added risk to that.

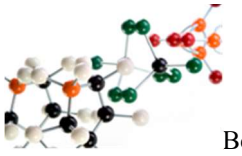
(17) **Question/comment:** What this is estimating is if the procedure kits go away which is part of the baseline number and we start having to open individual items, so this is the added risk to that. If it is added I don't see any additions. I should see some additions here or something.

***Response by Dr. Fraiser**:* It is in addition to the baseline rate.

(18) **Question/comment:** Yet all of these are numbers less than one, so how does this become larger than that?

***Response by Dr. Fraiser**:* Because you have got a couple of multipliers. Part of the equation relies on the number of device packages opened and in addition to that you have the number of surgeries or central line insertions that take place each year.

(19) **Question/comment:** Okay, I am having a lot of trouble following the units of the math. They are not really provided. You have already acknowledged that you have a major apples and oranges problem because even at the basic level of annual numbers of infections and the annual mortality associated with annual infections as compared to a lifetime cancer risk. That's a

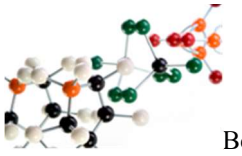


problem and not a problem that you created but associated with comparing infections and cancers. But I think you have to be much more explicit about the units and more explicit about the ultimate causal chain you're describing. You are describing the change from a multi-package environment to a single package environment, and the increased contamination risk associated with that. If you were to do all of this work on a per procedure basis, what is the incremental risk associated with the procedure with multiple packages on a per procedure basis, and how much ethylene oxide is required for each procedure? That's in two different situations. You are really talking about how much more ethylene oxide is ultimately being used or something like clarifying what has actually changed in the downstream consequences for ethylene oxide. What are the downstream consequences in the surgical suite? And when you are talking about multiple other types of infections you have not included, are those subject to the same issue of multiple packaging, or is the ventilator-associated risk now subject to the same issue etc.? So, I just think although I understand what you are trying to do and there's a long tradition of trying to do this, in the food safety literature and drinking water literature would be helpful there, but that's my general advice.

***Response by Dr. Fraiser**:* Thank you.

(20) **Question/comment:** We are working with a particular group in Georgia involved in collecting a lot of sampling and monitoring. And, going to a question by a panelist, there is a natural background of our body making ethylene gas and we convert 3% to ethylene oxide. There's a lot of natural background and urban level of about 0.2 µg/m3 and a rural background about half that (0.1 µg/m3) so it's in that range of 0.2 to 0.4 µg/m3. The particular group we are working with is fastidious about trying to knock this stuff out. When they did a reporting of their stack emissions, it was three pounds per year. And the whole point of the EPA assessment in calling out it being higher is a screening exercise so you could dig into exposure numbers and find out if we have a problem. It was a screening exercise. I think the Georgia EPD and the company looked at the exposure. They then got more accurate numbers which is the way it is supposed to work. There is this background level due in part to natural variation of ethylene gas, ripening fruit and that kind of stuff. Thank you.

(21) **Question/comment:** I hesitate to contribute much because this is not my area of expertise, but it's an extension of comment made earlier. I was thinking about bounding of the uncertainty about the assumptions that went into the estimates. I would be presenting the uncertainty bounds on those assumptions largely because they range widely. This informs sensitivity analysis of what is impacting most on the outcome of the assessment. Are the estimates potentially overlapping in relation to their very large bounds of uncertainty? And the other question I have related to these types of analysis is, whether or not alternatives have been considered? I am not familiar with the area. Are there alternatives to ethylene oxide that could have been considered? This, in addition to uncertainty bounds, would be important context, as a basis for interpretation of the outcome/impact of the analysis. You are comparing apples and oranges on one hand, but on the chemical side, what other alternatives do we have? It would be helpful to have the kind of perspective.



***Response by Dr. Fraiser**:* The FDA has issued a challenge. They are looking for ways to reduce ethylene oxide. There's a lot of work around reducing the amount of ethylene oxide being used to sterilize medical devices. I think they are having some success with that. There are some other sterilants that are being looked at. Already hydrogen peroxide is used, but the problem with hydrogen peroxide is you can't use that with anything that contains cellulose like Band-Aids or gauze. You can use hydrogen peroxide. It is effective but damages the material. I think they are also looking at gamma irradiation and a variety of things. People in the industry say there is no immediately available alternative that can be applied on this kind of commercial scale and not damage materials. I think also peroxy acetic acid is another one. There are some alternatives, and the other thing they talk about a lot and this is completely outside of my area of expertise, but to get an application approved, I think you have to go through a process at FDA and it's a lengthy process. I have heard people say it will be 10 years before we have an alternative that can be applied on this kind of broad, industrial scale that ethylene oxide is being used on.

(22) **Question/comment:** I am hesitant to comment for two reason. One, I have to recuse myself in the interest of transparency. It has to be noted that the EPA person is not endorsing any comments regarding which risk value is correct and that has to be clear. Secondly, I was wondering who is sponsoring your work? I may have missed that.

***Response by Dr. Fraiser**:* I did not mention but I'm doing work for a relatively small family-owned sterilization company called Midwestern Sterilization Corporation in Jackson, Missouri, and they also have a plant in El Paso, Texas.

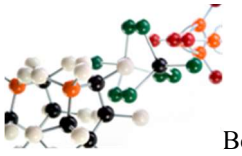
(23) **Question/comment:** I think that is important in any of these to acknowledge who is sponsoring the case study. I would echo the comment about engineering control. Was one of the context questions I also had an I would echo what was already said as well as not only a different type of sterilant, but different type of control technology. I think that would benefit this particular evaluation. And then, I also have a couple other questions on the equation as well. So, the second probability is microbial specific, but I did not see the infection was microbial specific. I was really hoping that another panelist can comment on whether that data is available because looking at the organisms considered, I would be more scared of a couple than several of the others. That seems to be an important variable. I was not clear whether it was addressed.

***Response by Dr. Fraiser**:* There were data on the fraction of specific infections like CLABSI or surgical site infections caused by certain bacteria. They vary widely across, but there were data on the fractions.

(24) **Question/comment:** I'm talking about the second subscript “inf”, should that not be subscript with specific microorganism? Was that considered?

***Response by Dr. Fraiser**:* Yes, it was. They are in the attachments.

(25) **Question/comment:** Yes. I don't remember seeing those. To get to that, the only data I was able to find – for example, what percentage of CLABSI are caused by Coagulase Negative



Staphylococcus? That's how I tried to address that. Of one of your other slides, you listed seven other organisms.

***Response by Dr. Fraiser**:* That's right. I was able to find data on five of them. There may be more data out there.

(26) **Question/comment:** I think it would benefit this to be explicit about which organisms were used and what the data sources were for the infection rates.

***Response by Dr. Frasier**:* Yes, I think I was specific in the paper.

(27) **Observer**: Thank you. I enjoyed the presentation. As a curiosity, my question is your earliest slide talked about the populations that can be at risk for people in the vicinity of the plant for ethylene oxide. And there is the person having the medical procedure done, and also a risk to opening up packages themselves. Is that right? Did you say that?

***Response by Dr. Frasier**:* There is a risk of contaminating medical devices by opening the packages.

(28) **Observer**: But the exposure to ethylene oxide from opening packages?

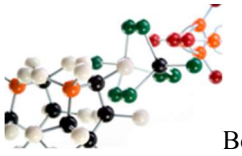
***Response by Dr. Frasier**:* I did not consider that. You could do that. I think there is actually an acceptable level published by the FDA.

(29) **Observer**: It may not be germane to the case study but the risk I would think would be higher for that person, nurses or technicians doing it many times a day, day after day, as opposed to the physician who does it and is done with that. Just curious, what you think the risk looks like? You presented cancer risk for people in the vicinity of the production plant but what is the risk for a person whose job is things they have today and the risk you show for people in the vicinity of the plant?

***Response by Dr. Frasier**:* I am reluctant to speculate on that because I have not looked at it. Part of the process is a lengthy period of off-gassing of the ethylene oxide from the devices and there is an allowable residual level of ethylene oxide published by the FDA. Some amount is there.

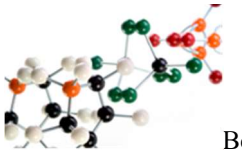
(30) **Observer**: So, are you talking outside the plant? And a quick question with the slide about the one paper about how often people change their gloves. That was from task to task on the same patient or from patient to patient?

***Response by Dr. Frasier**:* It was a percentage of glove changes between activities. I don't remember it saying whether it was for a single patient or from patient to patient. I would say for different activities with a single patient, probably most healthcare people don't change their gloves at all.



(31) **Observer**: We have almost exactly the same story as we just talked about. The chemical was classified as likely human carcinogen. It is similar in nature and also have extensive monitoring studies for ambient air. So obviously we would not pass the cancer risk so what we have done, and I started five years ago, the molecule has been undergoing cancer reclassification for the past 20 years. But recently we had a successful reclassification that changed it from likely to suggestive. It means we will not be able to use the current risk assessment method, but will use the current value using margin of exposure. From my experience, that's something that I'm thinking is: this might already be explored and look at it extensively, by looking at the current data set, is the tumor, the current tumor profile, does it deserve the current classification, which is one thing and the second is what is the tumor type? If it is vapor, then that would be usually the EPA dosimetry guidance. So, based on where the tumor is located, right? If it's local tumors like nasal, or pulmonary or systemic, they have different adjustment factors. Would you be able to do any chemical specific adjustment factors that can show that adjustment factor using as a default if it's more preservative and completely change your dose, so when you do the modeling you might have a different answer? That's one thing we have done. Secondly, 24 hours are not representative of the long-term exposure, and obviously not representative of lifetime exposure when you talk about the potential. It is exactly what we have done where we actually went to a greater level and developed a model exposure system so currently, they are evaluating the system. They want to use the system to evaluate the exposure in general.

(32) **Question/comment:** The ethylene dioxide cancer assessment is based on human epidemiology. We are talking about human cancer incidence. The controversy remains about how the EPA went about doing its risk assessment with respect to the implications of the number derived from the modeling of human epidemiology responses.



Case study 6: Risk Assessment Methods of Flavoring in E-Vapor Products

This case study was presented by Dr. Julia Hoeng, Philip Morris International, and Dr. Donna Smith, Altria. The presentation is available on the *ARA* website. The purpose of this case study is to is create a consistent and documented process to characterize chemical hazards, so that employers could make well informed decisions for those chemicals lacking OEL.

Discussion and Comments from Panelists and Audience

(1) **Question/comment:** I have a question that goes way back to the beginning. You may have said this, and I missed it. This is about the representativeness of the sentinel moiety of the flavoring group that you stabilized it. Correct? So, this is where I lost you. You have this moiety, it is one chemical within a group, a flavoring group, see you ended up 38 chemicals, right? And then, you stabilized it. Could you go through that again? I lost you at that point.

***Response by an author**:* Sorry, yes. So, the test formulation contained all the 38 flavor group representatives. Right, so, stay with me. That is what we would use in the actual application. The in vivo and *in vitro* data that you are looking at that they represented, that used the test formulation. But to make that, we found out of the 38 we group compounds that were similar to the activity and we structured those into pre-blinds which were concentrated, which were five to 20 times higher than the final formulation. Those concentrated pre-blends, we do not do any extra stabilization. It was just that those compounds, because they have similar structure and reactivity, that we hypothesized at the beginning they would be more stable. Then we showed data that the example of the preplanned that they were stable for weeks. It could have been even further, it only showed four weeks.

(2) **Question/comment:** Thanks, that is where I got confused. So, the moiety that is being used for any one of these chemicals, it would be the same form of that chemical as in the commercial product? Yes?

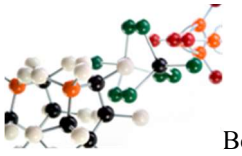
***Response by an author**:* This is a fundamental study so I want to be clear that these are not the same concentration levels that would be used in commercial products.

(3) **Question/comment:** If you have a flavoring chemical, the chemical pre-mixture, is this the same form of that that could be in the commercial product?

***Response by an author**:* Correct.

(4) **Question/comment:** That is where I had trouble tracking, thank you for that.

(5) **Question/comment:** Firstly, given the complexity of approach, here, there may be a few of these questions for clarification. I had a question about how you did the groupings of the different ingredients. In essence, how you set the bounds on those groupings.



***Response by an author**:* The recommendation of the European Community documents was that groups of chemicals, based on their metabolism and their moiety, are supposed to have similar biological behavior in common. So, we took these groups that are in this regulation. There they have 32 different groups. We have allocated these 250 chemicals based on the defining for the different groups. And, we've ended up with a 28 out of the groups. And, these 28 groups, we feel that there was not enough granularity to define the diversity of the flavors that we have allocated. That is why we have some groups in 8a and 8b in order to have more granularity, that is how we have to find the different groups.

(6) **Question/comment:** You started with something established, and was the established for flavors?

***Response by an author**:* Flavors for food.

(7) **Question/comment:** Just one more quick question. I'm curious as to why you chose ToxCast for the QSAR modeling.

***Response by an author**:* I found it was usable. It's tool for purpose. We found it was quite predictive. Based on previous experience. The same correlation for some chemicals.

(8) **Question/comment**: I will echo the compliment about how sophisticated and complicated your analyses are. I have a couple of questions to get the details on it and I may have missed it because we did go through quite a lot. First, you mentioned they were inhalable for in vivo testing. What are the MMAD and GSD of the device? Is it the same?

***Response by an author**:* It's very comparable. It goes up to 4 μm in commercial products.

(9) **Question/comment:** It does not depend on the device itself?

Response by an author: Depends on the device. And, for human beings, considered below 6 μm, or as for rodents, again, depending upon the recommendation, changes over time. It is below two or 3 μm.

(10) **Question/comment:** Okay. You had two different *in vitro* systems, one in 3T3 cells and the other, bronchial epithelial cells. Were the conditions the same? The assays were not the same. Were the conditions used the same?

***Response by an author**:* Similar. For the NRU, the treatment was four to 48 hours, only 24 for the RTC, which we used the primary cells, so the conditions are slightly different, and the cell system is quite different. We are talking about cell lines or primary cells. What is interesting is that, the vast majority of data were going the same direction. For instance, you can see the preplanned never classified, found to be the most cytotoxic, with the most cytotoxic with the two methods, despite the differences. There were some differences, especially when we assessed the full mixture. But, in general, the results were going in the same direction.

(11) **Question/comment:** When you say essentially similar, this is not an assay system but a merged culture system?

***Response by an author**:* Exactly.

(12) **Question/comment:** And, did you do any calculation of what the actual dosage was of the cells? The differences between these different types of cells give another opportunity for variability, it is a big debate right now in inhalation *in vitro* assays. I'd commend you on the open data platform that is the way to go.

***Response by an author**:* We have the systems in our facility and, for sure. Thank you for the comment. Actually, we do a lot of exposure studies, and we intend to actually ensure that we understand what is delivered to the subculture and how representative it is along the respiratory tract. We do take these things quite seriously. We haven't done it with this particular flavor formulation. On our to do list.

(13) **Question/comment:** I was curious. What was the dose tested for the 3T3 and the human bronchial epithelial? I don't remember. You go down to cytotoxicity. From there you divide it.

(14) **Question/comment:** My only other question on the plating was did you do tier or any other analysis to ensure it was confluent and the integrity of the epithelial barrier was established before you did the essay?

***Response by an author**:* Absolutely. When you say primary cells, we always do the tier first and the measurement. It's online determination of the toxicity of the cells. So, we only used concentrations that are, I mean, if you have more than 80% toxicity, we wouldn't use it.

(15) **Question/comment:** I'm not familiar with the Alexander correction. I was curious because I know the folks at Applied Research Associates, Inc. (ARA) that developed the Multiple-path Particle Dosimetry (MMPD) model, which takes into account the aerodynamics of the human respiratory tract. Why wasn't the MMPD model used to do dosimetry correction, especially if the end device particle size varies as much as you just said it did. I would like to see the dose correction done with the MPPD dosimetry model as opposed to that equation.

(16) **Question/comment:** We spend a lot of energy on extrapolation and we actually had a workshop year last talking about models for *in* vitro systems and I think we are developing models that consider the geometry of the exposure systems and different animal strains. In terms of gold standard, my agency considers the MMPD the standard, and we are looking into trying to do it for *in* vitro to *in vivo* extrapolation (IVIVE) also.

Response by an author: Exactly. We've contributed to the model with additional geometries. We actually just finished at the one geometry and were able to demonstrate that these have different geometries. The data will be implemented in the MMPD model. This is excellent for particles, for these types of evolving heirlooms. It's still a challenge but you have to understand the physics.

(17) **Question/comment:** I agree. You're probably aware that Mike Oldham and Owen Moss have shown strain differences in dosimetry. In fact, what an NIEHS researcher touted to be differences of genomics across strains can be described to differences across symmetry in the strains and they publish that years ago.

***Response by an author**:* They did in 1976, the doctor came back and said let's look at it for some of the strains we are using, it's a work that we are currently doing together. It's absolutely important. It may seem fundamental but it's important.

(18) **Question/comment**: Now for some really general questions. On page 3 you talked about AUL acceptable exposure or use levels and cut it. Sounded like a good thing to do, but you didn't further explain it or maybe I missed it. Could you explain more what is in AUL and how did you determine those, please?

***Response by an author**:* AUL stands for maximum use level and it's the concentration of a certain flavoring used in commercial products and if you put this one product in different types of products, the MUL would be the way you would put most of it. You have the same term for acceptable use levels for representative flavors. The acceptable use level would be what is acceptable as an outcome for this type of investigation.

(19) **Question/comment:** Okay, then the next question. You talked about something being lower on page 6. So furthermore, the flavor reaches 18% of the liquid mass. They account for much lower mass value and you said something about 6-fold lower though the actual use is roughly six-fold lower than 18%, so about 3%.

***Response by an author**:* Commercial liquid just uses between one and 3% of flavoring.

(20) **Question/comment:** Sorry, I am asking questions because this is great stuff. I noticed in your multiplex for the cytokines that you mentioned neutrophils are important, yet in the multiplex I didn't see the allarmins in the cytokines, could you consider adding alarmins when you do that? Because they are the key for airway epithelial cell (AEC)-mediated neutrophil influx and you are probably not capturing that.

***Response by an author**:* So, the cytokine panel use I hear is basically just a basic one you can purchase in a box from Luminex. We've been trying to find alarmins with mass spectrometry-based approaches but it's not trivial.

(21) **Question/comment:** Talk to Jack Harkema at Michigan State University.

***Response by an author**:* Thank you. Yes.

(22) **Question/comment:** First, I really commend you on the openness. It is what makes science better. I was just curious because you had limited time to provide detail, but are the results of the peer reviews available publicly as well?

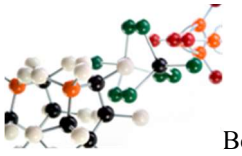
***Response by an author**:* The results of the peer review are available, and we wrote a publication on the summary of the entire peer review. I believe we had several experts to get all the data really ranging from pathologist to toxicologist inhalation experts, genomics experts and biostatisticians. It was an expensive exercise for Phillip Morris and not easy to find experts. For some people, they were really excited, and some said we don't want to look at this since this is from Phillip Morris International and that they would be getting gauged. But the reviewers are blinded to us, so we don't know who the reviewers were but clearly the people who got to see the data knew that it was us requesting the review. But it was not to say, hey, the product was amazing, but it was a feedback of the statistics. Are we looking at the right biology or are we missing something? So, we need to get the feedback.

(23) **Question/comment:** Just following up on the transparency and peer review concept. I know that obviously you have much more incentive to be transparent about a product that's likely to go across the finish line. For the one thing that I think a lot of the public do not understand, is how many things you would internally put aside and how many things don't come to the market because you have pre-assessed them as being unsafe. I'm wondering how that plays in your role because people often say here is the product, but they don't realize there are 200 other versions of that that didn't cross the finish line. The effort to describe that would be enormous but it seems to be what's missing. Do you understand my question?

***Response by an author**:* I understand your question, and it's very pertinent because even internally, it's taken us 10 years or more to come up with this product, which is now commercially available. So, 10 years of work of the organization and millions of dollars, billions that have been spent. So, yes.

(24) **Question/comment:** Just thinking it through. It's not that people need the level of detail because the level of detail is exactly why you rejected it and it's not as important and it's not exactly why something you're putting on the market is safe. But the fact that there are 200 variations of the thing in the 16 flavors that you didn't include or think of including because of whatever it was, you said to heck with that. Is there way of describing that or any precedent for doing that? Maybe, across the chemical industry is a broader question as well as broader question as well.

***Response by an author**:* I think you're making a really good point that we tend to focus on giving guidelines on how you should do something rather than saying, don't do this. But, clearly, the very biggest is the product offline, if you can't control the number of puffs or if the engineering is not enough, a lot of products use the technology that you can't manufacture in high-quality so, there are challenges in the design space. Then, the next part of the composition of the flavoring and of the solutions that you put in the devices that these are the areas where you have a toxicological concern and, we state that in our publications over and over and over. But we certainly don't pinpoint by putting things on the market that aren't safe. The only thing that we can do is basically describe what we do and why we do this in this way that is why it's close.



(25) **Question/comment:** Couldn't it be a simple workflow description? As Greg said, people don't want the detail, but they want to understand that you eliminated something before you got to this stage. But I don't think this has to be complex.

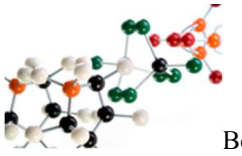
(26) **Panelist:** Yes, I think there are a couple of publications out there. I know Dr. Virunya Bhat has a publication. There are a couple of publications out there. I know Dr, Bhat has a publication against the last author that goes through basic product stewardship. If you're in a development process with iterations and what have you and you do, you start with chemistry. Because comparatively it's the cheaper thing to do but before you go out and do a full scale study that cost $10 million, you want to rule out those things ahead of time. So, most of the time when we approach it, we approach it from during the fastest and the cheapest and whittling things out. I think your point is a good one and that people don't get to see that a lot. So, I think, and I said to science is credit stewardship and thinking like flowcharts that you're describing, we look at this and we will spell that stuff out to minimize the biological impact it has over here. The more we can get that kind of information out, the better for the scientific community. That way, people can see that there's a thought process that goes into it and it's not just about the shining star that made it all the way through because that does look a bit self-serving.

(27) **Question/comment:** A good example, and this probably didn't happen but it's a good example of where it could happen. If in the early developmental processes of these devices you investigated the use of vehicles other than propylene glycol, like oils and vitamin E acetate, if that was discarded early on in the process as presenting health problems, we could've avoided a huge problem here in the United States, where with the advent of cannabis products, with those types of vehicles requiring an oil solubility vehicle that could've been immediately dismissed to say don't do that.

***Response by an author**:* Things like that, one of the things when you're looking at vehicles and carriers, it's one of the first things you look at and as someone who spent 20 years but someone has to use oil and figuring out that oils and lungs don't mix. We are coming out with a paper and using vitamin E acetate by looking at replicating the same findings in the animal models as we are seeing in the case also.

(28) **Question/comment:** Given the sophistication of the approach, here and the number of complementary evidence streams, I wondered if it would be possible to track the "learnings" in relation to designing such testing strategies with novel and traditional testing methods. Another aspect that would be important, relative to Greg's point on transparency about numbers of products which cross the finish line – perhaps an indication of % capture (i.e., proportion that go on for further investigation at any stage).

(29) **Question/comment:** The other question that I had, and I don't have a good recommendation or suggestion, but I wondered if you thought about it -- would be, given that we are using these plans, do we have any plans to try to describe what the principal component is that is causing toxicity in the particular different formulations?



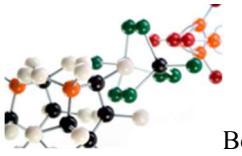
***Response by an author**:* In our experiments, we had 7 or 8 different flavors, but we have assessed chemicals for 240 mixed flavors. We have the data.

(30) **Question/comment:** How are you planning to compare that? Not only do we have an opportunity to look at this difference, with both *in vitro* and *in vivo* data, but you have a really awesome chemical mixture assessment in the data set here. I would encourage you to look at it that way as well. There may be ways to compare singular chemicals to what you are seeing with the mixture. I think that, you know, these are mixtures and risk assessors will be very keen to help with that.

***Response by an author**:* Thank you very much.

(31) **Question/comment:** I wanted to thank you for your presentation. It was one of the clearest and most detailed. I really enjoy seeing the energy and thought, particularly, and the strategy that you put into dealing with a very complex toxic logic scenario that you have presented. Fundamentally, I would say that your strategy looks like it is about as good as you could have ever imagined in terms of saying how can we go about getting to the heart of the mixtures that we are dealing with, minimizing the use of animals, and ultimately making decisions that are going to be right for public health with respect to this product. I really applaud that. Then, on top of that, on top of the strategy scenario, I think that you did something that was very important, recognizing that you are dealing with compounds that are inhaled, and as a toxicologist feels, data holds up significantly in toxicology. The strategy and ultimately the energy and thoughts that you had put into characterizing the dose, making sure that what you are delivering ultimately to the animals and your exposure was reflective of the product that will be out there, as a representative product. It was clear -- I have not seen many scenarios where this type of complex product is thought through to the level that you as a group did, in terms of making sure that whenever you might see in your whole animal model, you have optimized your opportunity for translation into really interpretable risk, as a consequence of the energy, the pre-energy that you put into making sure that your dose delivery is correct. A couple of observations that I had, and you are probably way ahead of this already, given where you have been. As I was looking at what you are doing in 28 day studies, for example, my experience has been with mice with chemical intoxicants that I have dealt with is that that is certainly fine, and it is not been a sensitive scenario in terms of detecting the range of toxicities that can happen in mice. The one tool that I did not see you using, maybe you are, and you did not tell us about it, it is using cell replication turnover. There are a cluster of cells, incredibly sensitive to damage. When they do, they go into a regenerative process. You will not see that physiologically if you are looking. Normally, the lungs will look perfectly normal, but all you have to do to detect the cell replication is to give them bromodeoxyuridine / 5-bromo-2'-deoxyuridine (BrDU). It is a good inexpensive tool that works, and it can show whether you are producing potential damage that is not cytotoxicity. It could be myogenic. It is another endpoint that is commonly used for mice lungs. It makes a tremendous difference in helping us understate the MOA and mouse lungs.

The second thing, that I would offer you to consider -- it sounded tome like you're already doing this, but again -- it ties into what we were just talking about a few moments ago. That is, using transcriptomics. That can translate to the cellular stuff. If you are doing *in vitro* studies, and you

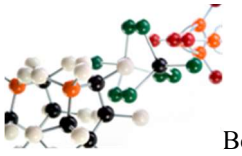


get a signal, or any one of those individual prototypical flavored inhalants, if you get a signal on the transcript, then you will ultimately go to whatever you see in the lung, whatever it will be. You potentially have an opportunity to narrow it down to which compound might be related to that effect, you will say aha, that is related to the signal that was produced by this 38 ones that you are now seeing in your cells. Transcriptomics can be -- it is also valuable, because we found out recently working with compounds like methylene chloride, and others that those transcripts on the date are incredibly useful for informing the mode of action in terms of potential implications for the endpoints that you are concerned with; long-term carcinogenic outcomes. It has turned out to be very powerful. In fact, in some cases, we were clearly able to identify doses that produced no transcriptomics signal at all, which is important, an important revelation. It is an incredibly valuable tool for understanding that there is not much happening in the organism as consequence of the exposure.

A third thing very quickly, another issue that you're interested in would be the potential for irritation in these products. You could take your mixture and do a 50-50 calculation, or put them in a plethysmograph, and see if they are getting respiratory depression. That is a one-day experiment, but it would very quickly tell you whether or not the mixture is producing irritation. Right now, you do not know if that is the case other than your pathological observations. It would give you physiological observations in terms of toxicity. That is on top of what I thought was a very well thought out strategy, implemented for a testing program from that.

***Response by an author**:* As a matter of fact, we have been looking into many of these already. For instance, we have done cigarette smoke, we understand the proliferation levels, with cigarette smoke and it did not change much. You know -- adding new flavorings may have activity, worthwhile looking into. In terms of transcriptomics, you may have seen that I did not elaborate. The system biology/toxicology part of the work is that we do collect lung tissue and tracheal origin tissue and do an analysis. Basically, we do this to understand that you can have changes like one or the other, one is going up or down. Having the transcriptomics, even if the change is not significant, you might understand the inflammatory process. The irritation part, when all of these products -- we have a well-established physiology platform. It enables us to measure respiratory depression. We have not done it with flavoring systems using an aerosol generator, but we are in the process of publishing this. It is similar to the book approach if we presented to rats, it may change something in the animal physiology or not. We have not seen any of the flavorings causing any sort of respiratory depression. The last point, I pointed to the fact that it is a good model to look into morbidities. We use the platform to assess respiratory physiology, looking into compliance-resistance in mice. As well as the pressure-volume curbs that we build to understand if there is a disruption of the proximal mortise compartment in the tissue. It has not always been applied to this; it is in the process.

(32) **Question/comment:** Working with mice, different from rats, I don't know if you have confronted this or not, but in the chemical solvent world that I work, and it is pretty typical that they would cause club cell injury. That particular MOA is highly likely to be more specific, not even human relevant. The reason for that of course is because the club cells have a particular tight cytochrome P450 stuck to it, which is nonspecific. The 2F2 isoform seems to have an affinity for compounds that are nonspecific; there is a class of chemistries that are reactive to it



too, but in humans, the human form 2F1 is inactive. You may confront a situation with your mouse. If you see damage to the club cells, we are going to have to do another iteration because it may be a real finding, but ultimately not human relevant. That is just the nature of the response.

***Response by an author**:* Thank you very much.

(33) **Question/comment:** I was excited that you had just mentioned that you are separating out nasal epithelia and epithelia from the lower respiratory tract, and thinking in terms of the aldehydes that you have in your mixes, and what is available in terms of dosimetry and MOA for humans, of the particular compounds. I would further encourage you to separate out the respiratory epithelium from the olfactory, because the cells have very different capacities, and that would -- as I was speaking to you at lunch, be how to do more quantitative IVIVE and this would facilitate it. If you can take this -- it is a big ask, to separate them. Their noses are small, but I have done it. If you can separate the lateral meatus from the septum, that would be more useful. It is a difference of both dosimetry and sensitivity. The lateral meatus airstream is 70 percent of the airflow for rodents. If you want to do a quantitative extrapolation to humans, you would need to be able to relate to what was actually delivered fluctuation-wise to the different regions. Both respiratory and olfactory epithelium and in both the lateral meatus and on the septum. You can see it, by eye, the difference in the epithelium. It is a matter of cracking the head open that is the hardest part.

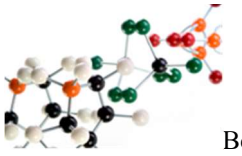
***Response by an author**:* I agree with your points, because, the rodents are nose breathers. The effects that you measure in the rodent nose are not the same as what you would measure in a human bronchial tree.

(34) **Question/comment:** We already have several computational fluid dynamic models for the chemicals in humans that would allow and facilitate a quantitative extrapolation.

(35) **Question/comment:** Thank you for your presentation, this is thought-provoking and interesting. I have a clarification question first, in the case study documents that you sent out in the discussion of your *in vitro* testing. You mentioned that at certain points in your studies, you were testing a full mixture of 38 compounds, and then at some point you were testing individual submixtures, and, the pre-mixtures. At some point you were testing individual compounds. I was wondering if you could clarify when you were testing those separate compounds, and why?

***Response by an author**:* What we want to see with this test is if we have high toxicity. So, we know that because of the clustering approach that we have applied, in the first step we said we have pre-blinds, and subgroups of flavor group representatives. We can test the subgroups. The flavors are the major contributors to the toxicity. That is why we have the pre blinds. And we tested the toxicity of components one by one. There, we were able to say general toxicity observed in the mixture is linked to these five chemicals. Is that appearing to be the driver?

(36) **Question/comment:** Okay, thank you. Okay, so this is more of a bigger picture question that is somewhat related. This is about your flavor toolbox approach. You know, in your approach you

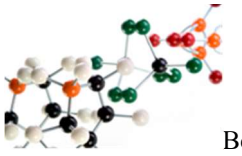


have done a nice job of breaking out different structural groups and selecting what is predicted to be the most toxic compounds from each of the groups. Then you put it all together in a flavor mixture. If you are thinking about this mixture that you have created, and how it applies, or how it is relevant to products that are commercial, for example, ultimately that is what is all this data is generated to try to inform on. You know, I wonder, if because you are selecting the most toxic compounds from each of the groups and putting it all together, maybe you are creating a mixture that is more toxic, or likely to be more toxic than what would find on the marketplace. How do the studies that you are doing with this relatively toxic mixture inform risk assessment or decision-making or anything else for commercial products?

***Response by an author**:* You're absolutely right. It was to create a representative mixture, to take the worst-case scenario that you could ever have. It is much more toxic than what you can find on the market. The aim here was let's create the mixture and see what we will have in terms of biological responses in the trial, and again, there were safety levels for chronic studies and this nature, we can identify a level, let's say -- corresponding to that. We can fix the level of this representative deaths, and this can be used to bridge the data and structure related components. Using the most toxic of each group, you can assume that with the endpoints that we have measured, or that were predicted, you can assume that the others from the same group left lower biological activity. You can actually apply the knowledge that you acquire on the flavor group representatives.

(37) **Question/comment:** You seem to be referring to the acceptable use levels that Mike was asking about earlier. Okay. You propose determining the acceptable use levels per flavor category, flavor compound? Can you elaborate more on those and how you will expect them to be used?

***Response by an author**:* Basically, for the studies, say we would identify a dose, or a concentration, in which we would see adverse effects. Once we have identified those, they could be low, medium, or high. One of the three that we tested. Let's say we went low. We know the concentration of each flavor representative in that mixture. This becomes the concentration for each flavor representative. Then this can be used in order to define the maximum allowed levels for deaths representatives, and then you can perform analysis within the specific group, keeping the flavor representative have restaurants, kind of. Do you see? This should be informed of intentional *in vitro* studies. We need to confirm that the flavor group representative that we have selected is actually the worst-case scenario for that group. So far, because of the lack of information, we use a lot of predictive tools. We will integrate with more studies, *in vitro*. We can use 3-D counters and other stuff to compare within the group the flavorings, and then when you confirm that they are going the same direction in terms of biological responses, then you can translate the maximum use level defined for the incident in the other member of the group. So far, because of the lack of information, we use a lot of predictive tools. In terms of testing flavors that were not part of our group, tomorrow we will do something new. You can run it through this as well once it is validated. Then, you can feed into the flavor group representatives, basically. If one is more toxic based on predictive data, or prediction data as well as *in vitro* data, then we will need to change the flavor representative. The new one would be the one that becomes the flavor group representative in terms of being used in commercial products. I am not the expert, but it is an approach where you have an index chemical and you can.?



(38) **Question/comment:** It is quite consistent with how we deal with combined exposure. I wanted to follow on the questioning. It is related to the question I asked originally, about the bounds of the categories. Are you interpolating and/or extrapolating toxicity data within the groupings of chemicals that you're using? If I understand correctly, the category is being formed principally based on chemical structure. You are determining the most toxic within the category (based on the testing battery and/or other?), and reading across to other category members, for groups defined elsewhere. Is that it?

***Response by an author**:* The group representative is based on two things, experimental data, as well as the data that is available -- we put the level to read across, based on the use level as well. The group gave an example, for instance, we use the maximum. The most toxic in the group. We would test at 1500 ppm to allow it to read across, within the functionality that determines the group, and in terms of concentration it has the highest concentration in the group. We will make a better job and presenting it.

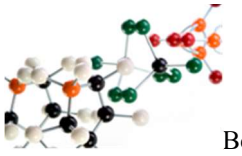
(39) **Question/comment:** As the only microbiologist on this panel, when I heard the word transcriptomics, I thought oh-- I am used to thinking about oral chemicals that encounter microbes before they would encounter the host cells in the stomach, but the same could be true in the lung, especially in the nasal cavity. Are you finding any in the extraction method that you used, you get microbial DNA and or the host DNA? Have you studied it at all? How the fragrances will affect the microbiota?

***Response by an author**:* I would love to. We have more ideas than we do space and the lab. You prioritize it a little bit, and we start to collect stool samples from the animal. We start to investigate the micro biomes to see if we see a difference in smoke. Do we see any difference? We do. That was already exciting. Now, we are taking this step by step. You need to expose the animal at least for three months, in order to start to see a difference in the microbiota. And the longer-term or chronic exposure studies, the microbiome is shifting. For ease of collection, we started off with stool, but, out of the scope from today -- we have animal models with inflammatory bowel disease, we expose them to cigarette smoke, and in the context, we do lots of microbiota work. It is a fascinating area.

(40) **Question/comment:** For the longer-term study that is coming up, will you look at microbiota issues for the longer-term exposures?

***Response by an author**:* Yes. One of the big challenges in the microbiota is not only do you extract the biological material, but what is the coverage or the sequencing? How do you do the computational analysis? We have a crowdsourcing platform, optimizing algorithms. We have a challenge open in a context of what is the right alignment method to find usable microbes that we care about. It is called SBB improver.

***Response by an author**:* Just to show you these groups, this is the European regulation, basically, they do find this is 34 chemical groups. You can see the description of the flavorings that should be allocated to the groups, and these flavors are supposed to have metabolic



considerations and biological and metabolic behavior in common. Like group 1, for us, that was too general. We added more granularity. We have allocated all the flavors into 250, using descriptions. It is the backbone of the evaluation approach. When we have allocated these structural and metabolic concentrations, this has been done on all flavors.

(41) **Question/comment:** My question relates to how you are aligning the members of the categories. In terms of reading across, for example, how are you aligning them? For a structurally related class. Are you aligning them based on comparative structures?

***Response by an author**:* It is according to toxicity.

(42) **Question/comment:** All right, grouping is by structure, but you only align them on the basis of toxicity?

Response by an author: Exactly. We split categories and this facilitates read across, basically.

(43) **Question/comment:** You are assuming for each member of the category that it is similar to the most toxic member of the category. I was thinking whether or not you could extend that by aligning chemicals within the category based on their biological profiles. This would provide a measure of relative toxicity within the grouping -- this is normally how we consider chemical categories, and combined exposures rather than considering all members to have toxicity similar to that of the most toxic component. I was wondering whether it would be worthwhile to do more here.

***Response by an author**:* The idea is to further facilitate this type of approach. This will happen two generations of *in vitro* data. We will have the other study where we can crosscheck the approach. In the end, the game is to rely on the one predicted status, as well as the validation data and say that we have built with *in vitro* biology.

(44) **Question/comment:** Part of the question relates to the fact that because you are doing so much work here, you have the potential to inform many evolving areas. In relation to combined exposure assessments, your observations could be extremely helpful in informing the development of structurally and biologically related categories for testing.

(45) **Question/comment:** Along the lines that we have been discussing, the research that you have presented today, using this constructed noncommercial flavor mixture, and also using a noncommercial capillary aerosol generator -- it seems to me, and I am speculating, tell me if I am on the right track -- it seems to me like the aim here is to sort of generate a bunch of data that is more applicable to e-cigarettes as a category. I am curious to know how that fits in with your broader research portfolio, because, -- at some point we want to know information about specific products, not just to the entire very diverse category. I am wondering how this work will inform or give information that is relevant to specific products that are out there.

***Response by an author**:* You are exactly right. This is category specific. This is not in any way product specific. The backbone of this is for the lack of the better term, qualifying the flavors,

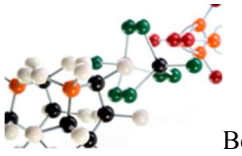
determining the flavors and inherent toxicity with these flavors compared to inherent toxicity of a solution without the flavors. We are coming up with a dose and being able to take that and extrapolate an appropriate use level for each and every flavor. That is what this is about. In terms of what you would have to put together, you keep talking about commercial and noncommercial, that is what you are getting at. The manufacturer would have to do work that was product specific. You would have to be able to demonstrate that the aerosol that comes out of your product falls within the parameters of an aerosol. You would have to look at chemistry. You can think of as a lot of chemical bridging. There is a lot of product specific information about paper products that lies and the device itself, the way it is regulated and the way the coil is wound, the metals that are in there, the gasket, the seals; it is extremely complicated. This is comparatively simple to what you would be dealing with in a product. I know that is scary, but it is true. This is comparatively simple, and, to throw phrase out there that makes all toxicologists in the world collectively cringe, something like this you could almost think of is inhalation graphs. You do all of the biological research, and you have the underlying systemic information. The countdowns will have data that have not been grandfathered. You have the data already and now you are going to have local specific effects and lung specific effects. You will be able to add to the scientific information that is out there, and then you will be able to pare down which compounds you can use and what the levels are that are suitable in order to introduce the amount of risk to your product.

***Contribution by another author**:* For the time being it is more focused on product development. Once it is more validated using the *in vitro* process, then, time will be right to speak about abridging. Then we will reduce the testing on animals. Then perhaps we will expand to a battery of *in vitro* tests, in order to make a toxicological assessment.

(46) **Question/comment:** Okay. To confirm one thing that I thought I had heard, it sounds like there is potential for this information to be used in informing product formulation then, as well?

***Response by an author**:* Absolutely, yes.

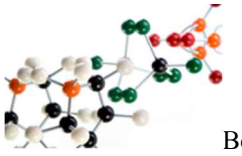
(47) **Question/comment:** So, I am not sure how to broach the subject, but I heard reference to effect levels. So, your testing and assessment strategy appropriately assimilates and generates data at different levels of biological organization (e.g., including transcriptomic). Normally, we compare dose-response across different levels of biological organization based on benchmark doses. On the other hand, you are looking at relatively non-toxic entities which complicate adequately characterizing dose-response curve as a basis for development of benchmark doses *in vivo*. As you move forward, how do you envisage characterizing dose response? I also think that you're assimilating a wealth of information that informs a number of evolving areas. I think that you made reference to one of them related to how we are going to capture this and have less animal testing and more *in vitro* informed decisions. This is a poster child case for these kinds of learnings. I would really encourage you to consider writing it up as a case study, in integrated approach to testing and assessment (IATA) It captures a number of important aspects - developing the test strategy as you become more informed and considering the value of data on hazard at several levels of biological organization to address combined exposures or mixtures.



***Response by an author**:* Thank you.

(48) **Question/comment:** Thank you very much for bringing the case study. I like the fact that you guys published at least one paper from a case study that was brought in last year, right? That's a good thing. I will pick up on a lot of responses that have been mentioned, from the big picture ideas. What people are most concerned with, I think, is, we are going to put chemicals in our lungs, because of the nicotine desire, and then -- will I get cancer from that? That is one thing. Is a level of the flavoring that you are going to put in safe? This goes to a point made earlier. So, some of your defense -- defense is the wrong word, but -- your rejoinder to that is generally regarded as safe (GRAS). It helps a little bit that you have recognized that we are talking about GRAS and inhalation. There is a difficulty there. One thing that you might do for the cancer question, is -- it is not an adaptive fix, because you are using mouse lungs. Mice develop the adenomas from the club cells which are decidedly different than the human disease. The human disease is a bronchial carcinoma, and it is not the same thing. In a 5 week study, you would actually be able to monitor lung tumor development by using certain fixatives, and doing a saline heart injection to clear the lungs of blood. With an ocular, you would be able to observe tumors that are to the level of 0.1 mm. So, if you want to do that, since you are already testing animal lungs anyhow, you would be able to do the observation. Now, again, that is not a direct and applicable case to the human disease, but it allows you to answer a cancer question at least in mouse lungs. You would be able to make a statement along those lines. The second thing goes back into my confusion over this idea of maximum use level or acceptable use level. To me, that was the safe dose. Right? The safe concentration. Data driving and extrapolation is the way to go, but, you are doing a surface area conversion from the mouse to a human, which is helpful, because we are all going to use that as one way to go from experimental animal tests to human. There is a toxicodynamic difference, and if we have data on this difference, we don't divide by a default factor of 3. The point is that you need to think about that perhaps. You have it within human variability that we always address with uncertainty factors, but we are trying to get away from defaults of 10 (or subunits of 3). Some applicable at the that you have there is that you have a mouth and the product that is 6-fold less than what you are testing. You have some idea there of coverage, but it has not called it out that way. Well, anyhow, that is the idea of this maximum use level, the acceptable use level. If you could just clarify that and address the fact that you are going from animals to humans, and you need to answer the question as well, you have this and the product, is it safe for me? That is what people want to know. We have a way to help understand that.

***Response by an author**:* Thank you very much. Maybe, a comment on the safety factor, so -- there are different levels of safety assessments. Yes, we use the surface area conversion, which is already a systemically cleared product. If we focus on lung toxicity only, because the body to body surface or lung to lung surface conversion is totally different, you go from a mouse's 0.1 square meter, to a square meter of the lung surface area in humans. You have an additional factor of safety, and obviously we have to concentrate -- the concentration, the final mixture, which is 80 percent flavoring, we have the fact that we have selected the most toxic ones and the mixture. If you look specifically for carcinogens, it is shown to be extremely susceptible. So, altogether, it brings already a certain level of safety. Is it 100 percent? I do not know.



(49) **Question/comment:** I am not trying to say I want to see a safe dose. I am saying -- you have uncertainties in the data. Of course, there are ways to approach a safety issue, and there are things built into your assessment that already do that. Clarify the terms, again, because I was confused. Maybe I misread it, that is all.

(50) **Question/comment:** We are probably not at that stage at the moment. You are making the assumptions that you are developing testing protocol, and, I also think that the computational fluid dynamic modeling will reduce uncertainty considerably as well. If any of the modeling can be done, that would really be helpful. I want to discuss the doses. You are talking about long-term effects. As you move on to do the studies here, this is a dilemma as well. We have regulatory guidelines studies, and we should be able to use much smaller group sizes and benchmark doses. So, again, I think that is something that we need to take into consideration particularly when you are looking at different permutations and combinations. If you want to develop a more product specific testing protocol -- so, I do not have the solution, again, it is more the question of negotiating with the powers that be in terms of traditional models for hazard identification or modifying standard study design to answer the questions that you need addressed? That is the dilemma.

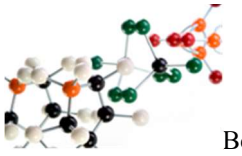
(51) **Question/comment:** I have another question, about the binning of certain types of flavors, into the categories, the toxicity on which -- when you choose the fentanyl product, the marker for the group, was that based on food flavoring toxicity studies? Remind me how you had determined that you had picked the most toxic one? The second part of my question is, if you look across the category, what is the range of the difference in the toxicity that you saw? How different was the potency process? Did it influence the way you put together your mixture?

***Response by an author**:* Okay, so, for the more toxic one, what we have done -- let's take one group. For all of flavors within the group, we collect LD50s. Then, we also collect unpredicted data like Cramer Class, then, within the group, we say okay, which flavor has the lowest metric for the different test point? So, then, we assign a number. Right? Okay. Let's say, we have 10 players. The lowest LD50 gets the highest score. Okay? After you accumulate all of these numbers together, and you have a score for the most toxic flavor within each group, that is how you aggregate all of these values, and you have a final ranking.

(52) **Question/comment:** Do you have data for all of the chemicals?

***Response by an author**:* This is the problem. Actually, experimental data, no, that is why we had to use predictive data, because with the predictive data, we can get a number for each flavor. Then we can do the final ranking.

(53) **Question/comment:** Okay. Before you do the chronic testing, or you look at this, let me ask -- from group to group, was there very much variation -- no matter how you got the score, how big was the variation?

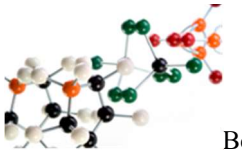


***Response by an author**:* We did not test the 38 groups between each other yet. But we have compared the preblends containing different types of flavor group representatives. Group A against group 1or a group to, or group 3, we are saying what kind of differences there are.

(54) **Question/comment:** Okay. On that point, before you do a long-term study, I can see where this would be an issue. Because, you are going to make a huge financial commitment and the long-term study, and, I would suggest that prior to doing so, you would have to revisit this idea of what is the mixture based on toxicity, because, not that I intuitively know what the answer should be, but the answer of creating a ranking score is terrific. That can be challenged. Even though you may apply it consistently, you are open to it being arbitrary, self-serving, you know -- how do you know that some form of one being a compound within a category that may or may not have data, when inhaled, might be uncharacteristically more toxic and should have been the seminal chemical for that group, before embarking on a mixture. Your question will get compounded when you mix them together, and then you will be doing the testing. I would recommend that you revisit this very basic first step and maybe you would be able to gather some good conversational pieces. I know that Jim does this kind of thing where you look at the potential toxicity of different things depending on the structures -- really, I know that you went through some of this. I would really emphasize that you should not have someone point back to the very beginning concept and say that that is not the mixture that you should have used. Also, someday if someone comes out with the toxicity data on one of these flavors, and you did not included in the mixture, then I would say that you are vulnerable, if you will, to an argument that you had had the wrong mixture in all of the tests. You might want to take a look at that, and even make more than one cocktail. If you are not sure about something, or whatever, I think that that is a vulnerable spot.

***Response by an author**:* We have a program to validate this approach and implementation. We have already run tests on all of them. We have screening going on. We have a talks tracker for 80 components already. So, we are building this evidence and we can probe the ranking based on the real data that is coming out of the system. That is part of the validation piece of this approach. But we had to start somewhere, obviously. So, where data was missing, we had simply *in vitro* tests with predicted data and as time runs, we populate this with more information.

(55) **Question/comment:** A quick clarification question. I only have one thing. Sorry. Say, given all of that and we do all that work that Patrick was talking about and everything comes out, we picked the right one and everything falls in line like it is supposed to, and let's say somebody publishes an inhalation study somewhere on that one particular chemical. And even though it didn't line up in an *in vitro* assay, you see the biological responses you didn't anticipate. From our company point of view, we continue to look on the interaction when things come up on things we use. As a matter of practice, that particular data, would it drive product reformulation? It would inform product development. So, that information would also be interpolated back into the process once the data came out.



(56) **Question/comment:** Very quick clarification question. Going back to what you are talking about in the beginning. You started with LD50 and some eliminated data on these compounds. They were from, I am guessing, oral ingestion, not inhalation? Okay, thank you.

(57) **Question/comment:** Yes. I think probably what this whole line of questioning speaks to is clarity on that point. And, from the perspective of thinking about it, because we have some experience in doing something very similar for 23,000 chemicals. And, we had a lot of data gaps. And, we had ways we weighted different information that we were using based on our knowledge of what they represented. So, I think probably a helpful way to communicate this was what was the minimum data that assured your comfort level? Again, I don't know these particular chemicals.

***Response by an author**:* I think we just need to do a better job in outlining what we screened for individuals as well as the systems used and what the mixtures were to present the results.

(58) **Question/comment:** Perhaps, indicate some criteria give you some comfort level that you had covered the area sufficiently to move on to the next one, kind of thing. That's probably what might be missing.

***Response by an author**:* That is a good point.

(59) **Question/comment:** Dr. Meek was referring to the Canadian domestic substances list that she and her team did about 12 years ago, I think. And it was one of the first classification systems trying to deal with massive amounts of chemicals, at least in the North America area. I am not sure where yours was at the time. I didn't follow the Europeans so closely. From this point of view, the Canadian work predated all of the work in California, and all of the work our group and others were doing. So, that is what the 23,000 refers to. Sorry, but I had to let folks know.

(60) **Question/comment:** So, I want to return to your 5-week animal study and ask if you could provide more details on how the exposure system was set up and how the exposures were done. You mentioned there were some errors that were generated from the system. And then it goes into various chambers and things. Can you provide some more information on that?

***Response by an author**:* Yes. So, let me start from the generation side. So, they had a fixed temperature which was set at 250 degrees. That is controlled. It is then diluted in the first dilution step to obtain a stable aerosol. If we dilute to the level that the animals receive right away, it would evaporate way too much water in the system; shows a dilution error consisting of 60% humidity. So, that is why we have different dilution steps. It goes from about five meters to enter the exposed chamber, which is basically a box with a sliding door that can be opened. The arrow is distributed to the back of the chamber through different tubings and we have a rack inside where we put in the mouse cages. That can accommodate up to nine mice per cage. The cage contains bedding material, but the performance verification of the system is measured with animals and with bedding materials inside. We know for instance if you have bedding material, you calculate the nicotine concentration and you know the yield to bring into the inhalation chamber. Without bedding material, you get a recovery of about 80% of your nicotine. Right?

Taking all the calculations into account, from the moment you have animals for and with the huge surface area as well as bedding material, your yields drop to 60%. Again, it is a fairly conservative approach, because in terms of nicotine and a lot of flavoring components, there would be a position on the front of the animals. And there is grooming that takes off more of the dose. It is again a little bit of a safety factor that is added into this type of experimental approach. When it comes to aerosol characterization, measurements are performed in between the cages. So, we insert sample collection tubing, for instance, to collect measured flow rate. And another to collect what the animals are breathing. So, all of the basic measurements performed on the aerosol quality and exposure levels are not in the breathing zone of the animals, and not somewhere in the tubing area. We collect samples in tubings and other system to understand where we do have losses and or changes in aerosol quality, but that is more on the technical side of the performance verification.

(61) **Question/comment:** I am wondering from the aerosol, after it is generated, between that to the actual exposures of the animals, how many folds are the aerosol being diluted?

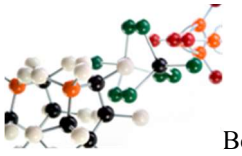
Response by an author: If I remember correctly for this study, we had a dilution factor of 376 from start to end, from the original generated aerosol and the output to the exposure of the test animals.

(62) **Question/comment:** So, the test atmosphere, the aerosol concentration, the nicotine concentration, how does that compare to what the person, for example, might get when they are vaping and how would you make that relevant?

Response by an author: So, we dilute to certain nicotine level in that atmosphere, right? Expressing microgram per liter for the specific study, 15 micrograms per liter. If you do the dose conversion based on pulmonary surface area, you end up with about 60 cigarettes per day for a human being weighing 60 kilos. That is not really realistic anymore, but that is the way it is.

Response by an author: So, there is a quantity of aerosol that the animals received obviously, in the whole conveying process of the aerosol to the chamber. The interaction with the fur of the animals you will have yields of different components. And, that is normal. You can measure a few by investigating biomarkers of exposure checking the urine. We also want you to understand that if you collect urine of a mouse, you get a few hundred micrograms. By the time you can sample it, it is dried up. You need to develop a lot of methods and controls, including some measurements to understand what quantity of urine you have two come to these values. It is quite complicated, but it is feasible. And, that is why for each of the pre-blends we have defined in order to make sure that we haven't forgot a pre-blend in the assembly of the final mixture, we will have one biomarker of exposure originating from substances of each of these three bands. So, the experimental setup is such that the nicotine concentration, for example, in the test atmosphere is at the predetermined level that you want?

Response by an author: Yes, 15 micrograms per liter. And by the way this is measured four times per day. So, it is impossible, and we are working on an online measurement of the nicotine



using infrared spectrometry. We are close to a final validation state. But still, we rely on the conventional approach where we trap the nicotine on the collection plate. In this atmosphere, we have sample periods of 15 minutes. And an hour and half later, we have to take it from the analytical labs. It is much easier because you have CO, which basically registers online and you can treat by the minute the concentrations you have in that atmosphere given the latency of this huge volume in the inhalation chamber. But it is easier than with these type of compounds. But again, feasible. And when we have the infrared spectrometry in place, we will be able to monitor in conjunction with one device of nicotine.

(63) **Question/comment:** Okay, so we have talked about nicotine and concentrations. What about aerosol characteristics? In the test atmosphere, have you characterized the droplet size and the aerosol concentration and things like that? And, how does that compare to commercial products?

***Response by an author**:* The characteristics we look at in the breathing zone of the animal is the particulate size distribution. It was shown in the slide. In our experimental system, with aerosol we are always depending on the composition of the liquid between one and two microliters.

(64) **Question/comment:** It was all measured? In the breathing zone?

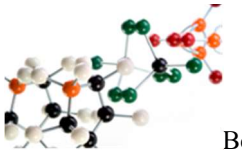
***Response by an author**:* Yes.

(65) **Question/comment:** Got it.

***Response by an author**:* And, we also know that it is no different from the diluted aerosol. Just as an information, when human beings puff on the system they will have the undiluted in the mouth, but with the inhalation piece there is a dilution factor of about 1-7 to 1-10. We know in these ranges that particulates size is not changing for this type of aerosol.

(66) **Question/comment:** So, something that you alluded to earlier is that in these types of test systems, especially when you have whole body exposure, there can be droplet deposition on fur and from that grooming, a little bit of oral exposure to these compounds. So, I am wondering, and I am partially asking the panel as well, are there better approaches that could be used that are still feasible to get around this issue?

(67) **Question/comment:** Shall I give her the short version? Obviously, there is something like a nose only inhalation system. If you think about carcinogenicity studies, the test guideline 453, you need to end the exposure with 50 animals per group, after the two-year period. So, knowing the mortality curve you need to start your studies, if you are speaking with much more than 120 animals per group. Once you do this response, you have the additional controls, both sexes turn up with easily more than 2000 animals on the study. In terms of practicability, exposing such huge cohorts of animals, even if you have an army of people is not tactical. What are the advantages? We do have data in one study and it is published, I think or it is in the course of being published where we compare the effects of no zone exposure to whole body exposure in one specific mouse train that we have used minus a mouse strain that we used for assessing



emphysema, as well as black lung formation. Given their genetics, they will develop some other conditions and we can measure this. We have seen that by putting these mice in tubes we have increased their stress level to such high degree that they don't even develop any plaque anymore. So, the feedback loop from the adrenal cortex is so huge in terms of anti-inflammatory pressure on the system that the whole inflammatory cascade that leads into emphysema as well as into vascular inflammation isn't there. It is always weighing what is the best for this specific model in terms of mode of exploration, number of animals, and expected biological outcome.

(68) **Question/comment:** Maybe just to add, from a human perspective, we have done collaborate work with Duke University where people inhaled these even deeper products and you can see it is a huge deposition in the gastric tract, in the stomach. We are very concerned about the animals grooming and licking, but we swallow it just as well as humans.

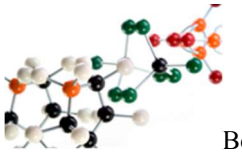
(69) **Question/comment:** It seems to me the issue of potential grooming induced dosing could be addressed in a short term. Since you are measuring urinary biomarkers of your nicotine, all you need to do is compare how those biomarkers compare in urine from inhalation, no zone only and if you see a trace amount more with the whole body, it gives you the formula right away in terms of how much is contributed by grooming. My guess is it is going to be so trivial it won't even be detectable in the variability of urinary test.

***Response by an author**:* For nicotine specific results, we end up with about 20% more biomarkers exposure in the same exposure later.

(70) **Question/comment:** Before I forget it, it is not related to your specific product per se, but it seems to me that the strategy you laid out has applications way beyond just the applicability to flavors and e-cigarettes. What I am thinking of is, there is brewing controversy that has no resolution. That is in the form of pesticides. There are many of them. Of the concept of formulations is the exact same as what you are doing. It is assumed that other than the active ingredient, the other formulas and the pesticides are essentially GRAS materials. That has been going on for the life of pesticide formulation development. But, there is mounting challenges that those formulations haven't been examined in perhaps to the degree we can, primarily because we have been challenged about how do we even think about that with all of the thousands of variations of potential formulations that are out there? It seems to me, you have just laid out a strategy of how that could be encountered. You can take the exact same strategy with the formulas used. Most of them are GRAS and you could do just a classification that you have done to start putting together, and perhaps we could begin to make a breakthrough in terms of that issue with respect to pesticide formulations. Using your particular strategy, it has some attractiveness to it beyond, and I think the creativity you put into this mixture issue has brought use way beyond the use of e-cigarettes.

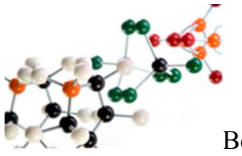
***Response by an author**:* Thank you.

(71) **Question/comment:** I would invite your comment or a little talk about this concept of presumed safety of this as compared to regular cigarette smoke. So, maybe you could speak to the whole idea of number (1): What is the number or range, or even types of chemical groups that are in the



e-cigarette that compares to a cigarette, not to mention the fact that you are inhaling the burnt products of the cigarette, as well is whatever you started out with? So, I remember on Tonga when we were looking at one of the additives on regular cigarettes, which by the way was chocolate, I had never actually considered chocolate was in a cigarette. From the get-go, it was a surprise to me. Could you speak to that a little bit? I know this is something you know so well. You may kind of jump over it a little bit, but for the rest of us who really aren't familiar with this whole industry, maybe you could speak to that and characterize why you think it is worth all this effort to develop these products.

***Response by an author**:* So, I will start and then I'm sure my colleagues will want to jump in. But I think it is well known, at least to us, that when you burn a cigarette you get somewhere on the order of 5000+ different chemicals. Commercial products in the United States, products on the market now don't have added flavors. They just have tobacco and water products. So, there are no flavors in those products. There are products on the market that have 50% market share and only have 19 ingredients, but to your point, things like cocoa, sugar are in the products. And there are products that have a lot more flavors and that. But the flavor load in a cigarette product is relatively low, much below 10%. So, it is very low. And, through the years when we were Phillip Morris and before we became Altria, we ran what was called the ingredient testing program. We published a lot of this is 2011 and Chris was the first author of inhalation talks. We buffeted through all 100 and some odd ingredients based on a ranking of how much of the actual additives were used. So, certain ingredients if they were only used at 11 bars per million, we would only run the *in vitro* assays. We did the chemistry because it is most sensitive, but least informative of biological activity. As you go along the spectrum, you may get less sensitivity as far as being able to determine differences. But it also gets relevant to being able to extrapolate bio logical data. What you basically find when you look at that across all the studies we ran is something we have come to recognize in the industry. That is something in the industry recognized as tobacco pick that is what drives it. What you see in the higher dose levels, most ingredients outside of experimental oils, that is a totally different thing and we won't talk about that. It has some unique effects. Actually, in the olfactory or in the eye. Olfactory, yes. We will set that aside for a minute because, we thought some of that in a primary component of Experiment oil. If you increase the level of ingredients upwards of 100,000 parts per million, in a number of those studies you start to see the toxicological endpoints ameliorate ever so slightly. Some of them are significant depending on which cocktails you add. We know at that point, if you continue to take tobacco out and replace it with the flavor, you don't see as high a toxicological response and you don't see any change in toxicological response at all. Heart of that to remember is, a good number of the flavoring chemicals in cigarettes are volatile. Because in order for you to actually be able to sense them, if you are going to inhale them, it has to get to the olfactory tract, which requires it to be volatile. A good number of those come off in process and are delivered as neat chemicals, very low levels for a lot of them. Some things like sugars are burned and therefore, they can produce other compounds, like these PAHs so, from all those studies, what we found is the flavors don't contribute substantially to the toxicology either one way or another on the whole. But there is that piece. If you go back to that being tobacco, when you look at e-vapor products, you no longer have tobacco. So, it is sort of a logically. If people are unable or unwilling to give up smoking cigarettes, and as long as cigarettes are legally marketed products, people are going to continue to smoke. Okay. So, if they don't give up



nicotine, what can you do to at least reduce the harm? So, that is where the e-v apor comes in. There is no tobacco, but it is just what you have in patches. Then you have the flavor ingredients which are carriers, which allow it to be aerosol and take it in. The flavors make it so people will actually use because nicotine by itself with carriers, you won't find a consumer out there who want to vape that. You are not. They are going to continue to smoke a cigarette. So, you look at the flavors and you like, this is what they need. So, the tobacco is not there, and we already know the disease in point. We know the risk associated with that. The problem with the e-vapor products is that we don't have any long-term data that we have on cigarettes. And, the literature is not really rich on the biological impact. That is part of what we're trying to bridge. Do you want to chime in?

***Contribution by another author**:* Yes. Thank you. We didn't show the chart today, but usually when we give presentations on these new product categories we say, smoking, and cigarette smoking we have somewhere between 5000 to 8000 substances. The next product class is a ≥90 reduction of the hormone constituents. In the e-vapor category, you have barely anything left you need to worry about. It is nicotine and a little bit of flavoring. So, the chemical space is dramatically reduced.

(72) **Question/comment:** Thank you. Again, a question from a totally different direction. As you look at your different categories and screen for activity, one of the big hot topics in toxicology of course, right now is endocrine specific responses. Have you been putting any thought into instead of screening for cytotoxicity as you do *in vitro*, as we do also screening for receptor driven toxicities of concern, possibly thyroid related? Is that factored into your conversations in terms of your screening programs for these flavorants in terms of whether or not they might possess those types of activities? Or, is that preselected out in those schematics that you used to characterize compounds into the different categories?

***Response by an author**:* That is a good point. We exclude endocrine receptors.

(73) **Question/comment:** How did you do that?

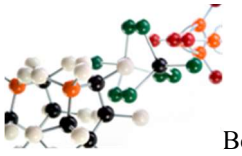
***Response by an author**:* Based on available chemical information. Yes.

(74) **Question/comment:** Structural, but not necessarily biological. So, in other words, you are looking for a potential structural characteristic of androgens or estrogens. Is that how you excluded?

***Response by an author**:* Yes.

(75) **Question/comment:** Ok.

***Response by an author**:* It is just based on the knowledge that is published, basically. If a product is reported as being similar with these properties, we just exclude it.



(76) **Question/comment:** I am just picking up a little bit on something I have been thinking about all along, but something you just said which was less exposure, less risk. The way it was characterized earlier, and I think you said this was a crude characterization, but you are basically making a list of possibly acceptable flavors. Right? So, it is prequalifying them for further consideration almost, right? Is there any thought to thinking about the pollutants of the flavor as a flavor agent, as opposed to its potency as a chemical? I'm sorry, as a toxicant? For the purposes of saying, I don't need to add very much of this at all to get exactly the effect I want.

***Response by an author**:* Believe it or not, that work is actually done in the product development organization, so there are people out there who do this for a living. I t is green. Here it comes. There it is. So, the product developers do that. To your point, there are some compounds that even those can be detected at extremely low levels.

(77) **Question/comment:** For instance, if you drop down to paper mill, everybody can smell the sulfur because the human nose can detect sulfur compounds. The same is true for some flavoring compounds. So, generally those guys don't want to add any more of anything than they have to. To get the desired sensory effect. It is not economic, and it makes the toxicologist's life hard, because there are times when we go back and ask, are you sure you need this much? Can you back it off a bit? Sometimes they are amenable to that and sometimes they can switch one compound out for another. Something I have learned over the years is there is more than one way to make a flavor taste a certain way. You can add certain flavor combinations to highlight what you are going for. Way out of my depth and expertise. I have just learned through osmosis over the years. That does go into it, just necessarily in our particular world. In other words, it would eventually be taken into consideration from this universe of flavors that are otherwise say, the only reason I mention this is that in the chemical alternative assessment business, which this is generally thought of when somebody has a chemical they are almost being forced to get rid of. Being forced to take it out, and the question is what are all the possible alternatives? For the longest time, that was purely a hazard-based discussion.

***Response by an author**:* Yes.

(78) **Question/comment:** And then, an assay report in 2014, it was basically said that there are inherent properties of this chemical that are purely exposure related and have nothing to do with toxicity. That should be part of the conversation and the inherent strengths of the flavor is inherently an exposure related variable.

***Response by an author**:* Correct.

(79) **Question/comment:** I was wondering if that would be something you put into the mix. Obviously, you are going to do it later because you have professionals who think about exactly that. It could be part of your safety assessment, as well.

***Response by an author**:* I agree. The easiest position to stand on is to go with your hazard assessment and be like, there is data out there on this chemical and we will take it out. That has sort of been the mantra over time. For my company, until about 2009 when, actually it was 2011

when we couldn't change anything. So you have to go through a different process. I think over time we have gone to more quantitatively based. To your point, there is a dose that makes the poison still in the route of exposure does, as well. I like to use water for people who can't seem to get that concept necessarily through their heads. But, agreed that over time we have been able to have more of these quantitative conversations about true risk versus just saying, that is bad and it doesn't matter how much of that thing you have in a product.

(80) **Question/comment:** Any kind of information on lung function tests in the people who have been using these products? Is that information available?

***Response by an author**:* The human perspective, not as far as I know for long-term use.

(81) **Question/comment:** And then the flavoring thing came, and I know when I used to be in the state of Texas, not all smells are good for everybody. What is nice for some is not for others. So, there is the threshold we used to look at. I don't know if flavors is the same thing.

***Response by an author**:* Yes. It is true, and that is why we have to provide consumers different types of combinations.

(82) **Question/comment:** I'm sure you know this, but I would just like to underscore it. The concept of FDA saying it is usable on foods or there is some magical list of GRAS or anything like that should be viewed, and especially if you are looking at the reputation of this kind of product, with some due diligence. Because, as you know and I am sure you know, even some of the GRAS products which is grandfathered in, it is like, oh well, we have already used those and we just put them on the list. Then, there were different periods of history across FDA's regulatory attitude, if you will, that these products got approved either with a little bit of data and sometimes a lot of data and sometimes with a panel declaring them or self declarations or whatever. So, the list, a lot of people think just because it is on the list that they must have gone through a risk assessment process. In fact, I saw last year the company appended because they made that assumption. And, more than half of the chemicals I looked at had not been tested. They had just been grandfathered in. And so, I mean you are obviously under a microscope with public scrutiny, and that would be a fatal mistake. So, the concept of relying on the idea that there has even been an inquiry or certainly that had any kind of risk assessment for chemicals approved for use on food, I just recommend you are very careful about exactly what that means when you say it out loud in public.

***Response by an author**:* Point taken. Thank you.

(83) **Observer**: I had a clarification question and a discussion question. Clarification question is about the use of the CAG system. I don't know if I missed something, but I wasn't clear as to why that system was used instead of conventional vape system, which would be basically the real thing. And, I was really curious about the slide where you showed some comparison information between a representative e-cigarette and a CAG system. It was mostly comparable, except for at the bottom there was something about formaldehyde. I didn't really understand how it went from that point, because it seemed like -- I couldn't tell because it was presented by

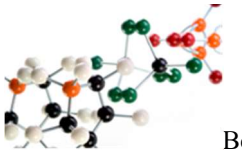
mass metric. But, from the actual format a hard formaldehyde content of vapor, I couldn't tell with the concentration basis we were talking about if it is like an 8-fold decrease in light of vanishing small amount of formaldehyde. But basically, I am not sure if I missed something, but basically why can't you use a conventional vape system? Which, I believe there are systems in which that can be used for animal experiments *in vitro* that people have been using. And then, are you concerned about the formaldehyde thing or am I missing something?

***Response by an author**:* The use of the system is practical. If you want to regenerate aerosol for inhalation chambers, we have two chambers connected for the high dose group. If you want to use e-cigarettes, you need to have at least four carousels with 30 e-cigarettes running in parallel to get there. These products need to be acquired and they need to be validated. It is much more work to get to a similar result. And your results will still be linked to a very specific product. We wanted to test a more generic approach that we could fully control, and we could fully characterize all the parameters and make sure we have a consistent result of the study. That is what I wanted to say about that. In terms of formaldehyde, if you compare it to conventional, yes even in e-cigarettes, in most e-cigarette there is a dramatic lower level than in cigarettes, except for the systems that do not have an end of cartridge system where basically you have a coil technology. The wick is in the fluids and if the fluid gets exhausted, the wick has less and less quantity of material and therefore, there is no control of the temperature of the power at which the coil would be heated up, in the end it will burn the residual portion of the carrier that is on the wick and generate formaldehyde.

***Contribution by another author**:* So, just to add. The aim of the experiment is not to represent the aerosol, what you would get from a device anywhere in the shop. But, the aim of the experiment is to represent an aerosol that comes out of a valve manufactured product that is fully characterized. So, that represents something that Philip Morris or Altria would be putting on the market. There is a lot of publications and some very recent applications that came out and you can look for the market map analysis in which you can see the aerosol chemistry from 50 different products. Usually, as you look at the chemistry of the last 50 puffs, formaldehyde levels really increase. A product like we are making in Philip Morris, you will not see that because they are very well manufactured. So, the aerosol that is generated is going to be representative of the aerosol you would obtain from a valve manufactured product. Okay? So, it was not intended to represent the worst in class. But this is the aerosol from the best in class.

(84) **Observer**: So, you are saying that if you had a better comparison experiment that represents vaping, where basically the system isn't being by the user?

***Response by the author**:* That is what the aerosol should look like. There are papers coming out where you have different groups in the world generate e-vapor exposures on animals and short-term exposure and they say the e-vapor product causes lung inflammation. Well, let's go back and understand the aerosol. How was it generated? All right? A lot of the signs comes up and says e-vapor products are bad for XYZ reasons. Let's look at what the aerosol was that was generated and what devices were used. Right? I am not thinking the science is bad, I'm just saying people need to pay attention. Usually we tend to jump right away to the "oh my God",



the biology. There is lung inflammation. What was the amount of aerosol actually given to the animals? Is it relevant?

(85) **Observer**: Thank you. That is where science and policy sometimes get mixed up.

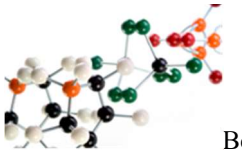
Response by an author*:* Absolutely. One more thing, we mentioned it briefly, but that was about those standards when it comes to testing batteries and how to do things. That also goes to parameters. A lot of studies that Julie was talking about, you have to look at the aerosol. You have to look at the puffing parameters they use. If you take puffs too long or too short together, or whatever the differences is and however they choose the puffing parameters are whether or not there is something close or something they made up, because we see a lot of that. That has a lot of influence to what Patrick was describing with the wick. If the liquid doesn't have time to re-saturate the wick, you are changing everything, the heat sink around it and how that device operates. So, that is one of the major needs is standardization around these puffing parameters.

(86) **Observer**: Thank you. So, a question I had for more discussion, here at NIOSH we have a lot of interest in developing sensory irritation as a human health and point. And, during the discussion a while ago it was mentioned that you did do some work with actual respiratory physiology, including respiratory depression in rodents, etc. That is good and getting to other kinds of endpoints that I tend to think about it. Your original workflow within silicone you and future phase and finally going to the animal, I did not appear to have any endpoints that reflect sensory irritation. I am not criticizing you for that. Heart of what we are interested in is developing better tests and developing better policy to make better use of data about that endpoint. And my question is, I was just curious if that has been part of your internal discussions, and if there is anything that you are looking at that could be incorporated into this work flow that might reflect a sensory irritation hazard that would be relevant to the airway? Whether it be something in the silicon phase where you, and I don't know how, you feel about looking at reactivity based on that, or if you are more of an adherent of conventional rodent based assays of respiratory depression? I am curious.

Response by an author*:* That is a good point. For the time being, we are not showing the data. We are performing some other tests. One big caveat is that you cannot appreciate the volumes in the correct manner, but you can appreciate frequency, which gives you a good impression over that action. If there is an irritation in the respiratory tract the frequency will get lower. As a consequence, usually animals compensate a bit with the volume, but the mid volume is usually lower. That is a given. So, we have the frequency in this test, we just didn't show it. It is fine. Obviously, we are open and constantly looking into additional types of tests. Especially with this category of products, where you get potential sanitization effects due to flavorings.

(87) **Observer**: We are looking into both *in vitro* and other assays to learn more there. There is one company now, I think they offer upper respiratory sensitizer assays, which we are trying out. They are not yet actuated or validated.

Response by an author*:* I was aware we were in the middle of figuring out the difference between the difference in terms of biological pathways between allergic and nonallergic



respiratory desensitization. I am always curious how groups that I don't usually talk to have been thinking as much about it as we have. Thank you.

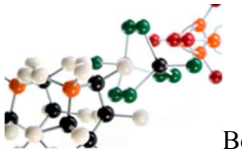
(88) **Observer**: Just one second please. These e-vapor products are for smokers, right? So, this is not something anybody should start consuming, right? They are intended for smokers who cannot quit otherwise. And that is kind of where I think she wanted to go.

***Response by an author**:* Yes, that is that the animals upon exposure don't have a choice. But humans do. So, something we have seen in studies that the soft and fuzzy behavioral scientists will do, things like actually use study and perception studies and behavior studies. If people can actually register the fact that this is irritating, or I don't like this; they are not going to use it. So, the self selection of those types of things that goes on with these types of products when people are looking for replacements. So, they have gotten used to the irritation of a cigarette. They expect that, believe it or not, people get used to it. But when it comes to a different product, they totally change their behavior. And it is a different thing they are using. They may not be as willing to accept some of the irritation. Just a little human part into that.

(89) **Observer**: I mean, it tastes good. You mentioned a few questions back that if it doesn't taste good, people will not use it, even if it is part of their goal of quitting smoking.

***Response by an author**:* That is the challenge. They have, the consumer has to like it and it has to meet whatever their needs are. And believe me when I tell you, each individual has different ones. And they have to find it satisfying. They have to like it. And, quite frankly, they have to make to want this twitch. It is all those behaviors linked together.

(90) **Observer**: Thank you.



Closing Remarks

By Dr. Michael Dourson

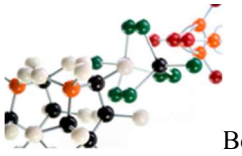
Okay, so if there isn't anything else on these case studies, we would like to close this part of the workshop. What I want to do now is to thank some people. First to thank are Heather King and Steve Gilbert of NIOSH, and Christian Williams and Bethany Hansen of TERA, all of whom have coordinated this whole affair from behind the scenes. Their support was wonderful. We also thank Frank Hearl of NIOSH for hosting this workshop. Christine Whittaker of NIOSH gave a nice keynote address, and, of course, John Snawder of NIOSH and Christine gave the first two case studies on occupation banding and real time exposure monitoring that were both well received. Sabina Lang and Lalita Shrestha of TCEQ did a case study on the use of epidemiology data that was helpful in understanding how improvements to these studies can eventually help risk assessment.

We then had a series of the five presentations, Bernard Gadagbui of TERA gave a talk on developmental toxicity and PFOA that was a case study at last year's *ARA* meeting; the paper is now published and has won an award. Lawrence Tannenbaum of the US Army Public Health Center gave a presentation on the truly adverse dose, which is a challenge to toxicologists by asking whether the critical effect is truly adverse or just a precursor, and what does that effect really mean to the animal. John Doherty (retired, EPA) and Dr. Carol Burns (retired, Dow) both gave talks on how to better use epidemiology data in risk assessment by suggesting novel approaches. Clare Thorp gave a talk on an integrated approach to modeling exposure data with toxicology data across different conditions and routes of exposure.

Our case studies then continued with Neeraja Erraguntla of ACC who gave one on ototoxicity from co-exposures to noise and chemicals involving multiple organizations in a collaborative fashion. Lucy Fraser of Lucy Fraiser Toxicology Consulting then gave a case study of the risk/risk comparison of ethylene oxide exposures and microbial contamination, which as we all know is a topic of current interest. Finally, Julia Hoeng of Philip Morris International and Donna Smith of Altria gave a case study on risk methods for flavorings in e-vapor products. This is a follow on to a case study they gave at last year's *ARA* meeting that has also been published.

As expected, this workshop generated a lot of good interaction. This is due in part to the Science Panel, including core members Chris Chaisson of LifeLine, Annie Jarabek of EPA, Bette Meek of the University of Ottawa, Greg Paoli of Risk Sciences, and James Bus of Exponent, along with our ad hoc members Carol Burns of Burns Epidemiology Consulting, Krystin Carlson of NIOSH, Peg Coleman of Coleman Scientific Consulting, and Cissy Li of FDA.

And finally, the Risk Assessment Advisory Committee, Neeraja Erraguntla, Pamela Williams, Sabine Lange, Suzanne Fitzpatrick, Mark Johnson and myself all welcome your comments on how to improve this workshop series.



Ongoing Activities

The presentations under "Ongoing Activities" are available on the ARA website.

1. **Data-derived Extrapolation Factors for Developmental Toxicity: A Preliminary Research Case Study with Perfluorooctanoate (PFOA)**
   Dr. Bernard Gadagbui, Toxicology Excellence for Risk Assessment (TERA)

This activity was presented last year, and the panel provided useful comments, which were addressed. A manuscript was submitted to a peer-review journal and was accepted and published. The purpose of this presentation is to give some highlights about what we did, why we did it, and what was the result. Based on additional evaluation, we determined that the average concentration during the various exposure windows of concern (i.e., during pregnancy), and not the Cmax, would be the appropriate dosimeter. A Data Derived Extrapolation Factor (DDEF) of 1.3 or 14 was determined for PFOA. These values are significantly different than comparable extrapolations by several other authorities based on differences in PFOA half-life among species.

2. **A Repair for Non-Cancer Assessment: Introducing the Truly Adverse Dose (the TAD)**
   Lawrence Tannenbaum, U.S. Army Public Health Center

Larry Tannenbaum gave a presentation on a concept he developed and recently published on - the Truly Adverse Dose (TAD). His talk noted that for a large percentage of noncarcinogens in IRIS, there is no information to demonstrate that the identified "critical effect" is harmful in any way. In that case, while a toxicological basis for deriving an RfD or an RfC is lacking, risk assessors continue to needlessly compute hazard indices for the many chemicals. Larry's presentation identified an intact animal dosing study approach that, if implemented, would identify doses of concern, should a given chemical have these.

3. **Proposal on the Epidemiology Peer Review Council**
   John Doherty, EPA (retired)

Epidemiology studies appearing in the open literature that have a mixed bag of GLP, quality assurance and ethical standards are often accepted at their face value by regulatory agencies. It is very difficult to obtain critical data for an independent analysis. In contrast, animal studies especially for pesticides are currently conducted under strict guidelines, GLP, quality assurance and ethical standards are reviewed by being rewritten into Data Evaluation Records (DERs), subjected to multi levels of primary and secondary review and further review by specialty committees and are subjected to on site audits. This disparity in the level of review is not serving the public well. Toward this end, I would like to propose two suggestions. The first, standardize the way animal studies are submitted to eliminate the need for transcribing into DERs. This will render more quality time to detect study deficiencies and subtle responses not already reported. The second concerns the formation of an *independent Interagency Epidemiology Peer Review Council (IEPRC)*. The IEPRC would consist of six independent subcommittee as: *i*, ethics (including means to obtain detailed data), *ii*, endpoint analysis

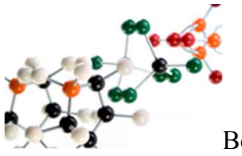
(consisting of experts on the endpoints reported), *iii,* statistics, *iv,* exposure, *v,* analytical chemistry and *vi,* animal toxicity and SAR. Each sub-committee will independently review the study in terms of their respective disciplines. The Council will then convene to evaluate the reports of the subcommittees and determine if the study supports a significant hazard for the chemical or otherwise. The Council's report would detail all justification for its decisions and objecting parties would need to specifically address these justifications. This process should demonstrate to the public that a thorough review of the study was conducted and hopefully reduce the need for litigation. More details on the duties of each subcommittee and the role of the Council chairperson will be presented at the meeting.

4. Novel and Integrated approaches to modeling aggregate exposure to chemicals across different conditions of use and routes of exposure
Clare Thorp, SVP Crème Global

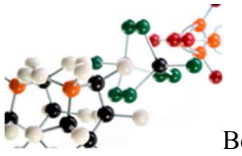
Exposure assessments are a critical component of a risk assessment. Existing models often rely on deterministic approaches, or simulate exposure data using conservative, default assumptions. Evaluating aggregate exposures across multiple uses, from multiple products and across different routes of exposure under realistic conditions is complex, and further complicated by the need for industry data which may be confidential or dispersed across many companies. This presentation will speak to how these data can be gathered, organized and pulled together within a probabilistic exposure model that uses Monte Carlo simulations of data distributions to calculate realistic aggregate exposures. The presentation includes a case study example of how Creme Global created the RIFM Exposure model with the collaboration of members of the Research Institute for Fragrance Materials. It provides an overview of how probabilistic modeling works in practice, and how it can generate data distributions that accurately reflect centiles of exposures as opposed to simple means. It demonstrates how capturing confidential business data and combining it with other existing datasets in the public domain, from the scientific literature or from other data providers can be modeled to address data gaps and generate the inputs required to model aggregate exposures in a refined and realistic way. Finally, it speaks to how the exposure data generated by these models can be combined with relevant toxicological reference doses or thresholds of toxicological concern to evaluate margins of exposure.

5. Bridging the Epidemiology Risk Assessment Gap: An NO2 Case Study of the Matrix
Carol Burns, Burns Epidemiology Consulting

A Matrix has recently been presented as a tool to assist in the translation of epidemiology literature for the needs of risk assessment. The Matrix is an approach to bridge the epidemiology-risk assessment gap and includes nine risk assessor "asks" of epidemiology studies. The Matrix is designed to facilitate awareness about how choices regarding a study's design, analyses and reporting can enhance the use of epidemiology data for risk assessment, and ultimately public health decision-making. This case study of the Matrix is presented to (i) demonstrate how a selected body of epidemiology literature can be described in a risk assessment context using the elements of the Matrix, and (ii) assess the clarity and utility of the Matrix. This was conducted by reviewing 14 epidemiologic studies used in the US National Ambient Air Quality



Standards Integrated Science Assessment for Oxides of Nitrogen for evaluating the nine elements outlined in the Matrix. The Matrix performed well in characterizing the needs of risk assessors in the areas of hazard identification, exposure assessment, and dose-response assessment. The case study revealed areas in which more precise word choices within the Matrix may improve the characterization of translational needs in epidemiology literature; recommendations for modifications to the Matrix are made. The case study findings indicate that there are opportunities for risk assessors and epidemiologists to collaborate to facilitate the use of epidemiology research for public health decision-making.



Appendix

Table of Contents

WORKSHOP INFORMATION AND SPONSORS **97**

BACKGROUND & PURPOSE **98**

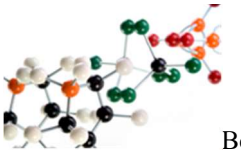
WORKSHOP OBJECTIVES **98**

LISTING OF RESEARCH CASE STUDIES **99**

COMMITTEES OF THE ALLIANCE FOR RISK ASSESSMENT **99**

FINAL WORKSHOP XI AGENDA & PURPOSE **101**

BIOGRAPHICAL SKETCHES OF WORKSHOP CO-CHAIRS, SPEAKERS, PRESENTERS, AND SCIENCE PANELISTS **104**



Workshop Information and Sponsors

Workshop XI Title Beyond Science and Decisions: From Problem Formulation to Comprehensive Risk Assessment

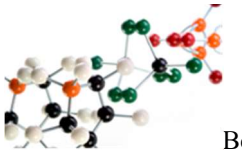
Workshop XI Site National Institute for Occupational Safety and Health
4676 Columbia Pkwy, Cincinnati, OH 45226

Workshop XI Dates February 18 - 20, 2020

Since 2010, a number of organizations have sponsored this workshop series through endorsement, in-kind donations, or grants. For a partial list of these sponsors see: https://tera.org/Alliance%20for%20Risk/ARA_Dose-Response_Sponsors.htm.

Current Sponsors/Endorsements for Workshop XI

- American Chemistry Council
- American Petroleum Institute
- Center for Food Safety and Applied Nutrition of the US Food and Drug Administration
- Consortium for Environmental Risk Management LLC (CERM)
- E Risk Sciences, LLP
- Exponent
- Georgia Pacific
- Gradient
- International Association of Plumbing and Mechanical Officials, Research and Testing
- The LifeLine Group
- Lucy Fraiser Toxicology Consulting LLC
- National Institute of Occupational Safety and Health
- Nickel Producers Environmental Research Association
- Summit Toxicology
- Texas Commission on Environmental Quality
- Toxicology Excellence for Risk Assessment
- US Army Public Health Center



Background & Purpose

Background

The Alliance for Risk Assessment (*ARA*) sponsors a series of workshops titled *Beyond Science & Decisions: From Problem Formulation to Comprehensive Risk Assessment*. Building on the ideas of the National Academy of Sciences' *Science & Decisions: Advancing Risk Assessment* (2009), nine workshops were conducted from 2010 to 2015 that brought together over 60 organizations seeking to clarify and advance the NAS recommendations (see: https://tera.org/Alliance%20for%20Risk/ARA_Dose-Response.htm). A total of 40 research case studies were presented at these workshops, which provided a real-time compendium of practical, problem-driven approaches for "fit for purpose" risk assessments. Specifically, the compendium links novel and evolving scientific methods and approaches with specific problems faced by risk assessors and risk managers in a variety of organizations (e.g., local, regional and federal governments, academia, private sector).

Purpose

Due to continued demand for the types of work products achieved by these workshops, the workshop series is continuing in 2019 and 2020 and will expand upon the discussion set forth by *Science and Decisions: Advancement of Risk Assessment* (NAS, 2009). These workshops will be conducted under the aegis of the Alliance for Risk Assessment (*ARA*), a broad-based coalition (see: https://tera.org/Alliance%20for%20Risk/index.htm).

Workshop Objectives

- Improve the risk assessment process by developing an updated and ongoing compendium of practical, problem-driven approaches for "fit for purpose" risk assessments, linking methods with specific problem formulations (e.g., prioritization, screening, and in-depth assessment) for use by risk assessors and managers at a variety of levels (e.g., states, regional managers, people in a variety of agencies, and in the private sector).
- Implement a multi-stakeholder approach to share information, ideas, and techniques in support of developing practical problem-driven risk assessment methods.
- Identify effective and meaningful problem formulation, and useful hazard identification, dose-response, exposure assessment, and risk characterization techniques for specific issues, including consideration of relevant data, description of assumptions, strengths, and limitations, and how the techniques address key considerations in risk assessment and decision-making. These techniques should appropriately reflect the relevant biology (including the biology of thresholds), mode of action information, and exposure variability at a level of appropriate detail.
- Provide methods to explicitly address human variability in assessments, including explicit consideration of underlying disease processes and exposure conditions, as appropriate for the relevant risk assessment context.

- Identify methods for calculating the probability of response for noncancer endpoints, as appropriate for the relevant risk assessment context.
- Identify useful decision-making approaches that incorporate risk information and uncertainty analysis.
- Develop a risk methods compendium that will serve as a resource for regulators and scientists on key considerations for applying selected dose-response or exposure assessment techniques for various problem formulations, with suggested techniques and resources.

Listing of Research Case Studies

The recommended framework for the workshops and research case studies is currently being restructured. For access to any of the prior research case studies, please see https://tera.org/Alliance%20for%20Risk/Workshop/Framework/ProblemFormulation.html, or contact Michael Dourson with Toxicology Excellence for Risk Assessment (TERA) at dourson@tera.org.

Committees of the Alliance for Risk Assessment

- The Alliance for Risk Assessment **Steering Committee** (SC) will provide guidance and oversight of the workshop series and research case study selection. The Steering Committee will have the final decision on charge questions after consultation with the Risk Assessment Advisory Committee and will have the final decision on members of the Expert Panel after a review of all nominations. The SC consists of state, tribal, and federal governments, academia, and environmental NGO:
    - Annette Dietz, Portland State University
    - Michael Dourson, Toxicology Excellence for Risk Assessment (TEAR)
    - Michael Honeycutt, Texas Commission on Environmental Quality (TCE)
    - Matthew McAtee, US Army
    - Moiz Mumtaz, Agency for Toxic Substance & Disease Registry (ATSDR)
    - Ralph Perona, Neptune & Company, Inc. [representing tribal interests]

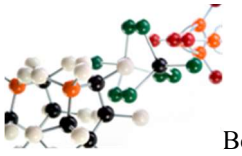
- The **Risk Assessment Advisory Committee** (RAAC) will be composed of state, federal, industry, and NGO representatives. This group will represent the various sponsors in the development of workshop structure, charge questions, development of Panel nominations, and the recruitment of presenters. The RAAC will have the final decision on workshop structure, presenters, and content, after consultation with the *ARA* Steering Committee. Members include:
    - James Bus, Exponent
    - Michael Dourson, TERA
    - Neeraja Erraguntia, ACC
    - Suzanne Fitzpatrick, US FDA
    - Mark S. Johnson, US ARMY
    - Sabine Lange, TCEQ

  - Pamela Williams, E Risk Sciences, LLP

- The Beyond Science and Decisions **Science Panel** (SP) provides input on research case study methods being proposed to enhance the risk framework. Panel members also provide input on the utility of the research case study methods to address specific problem formulations and identify areas for additional development of the research case study and/or method. Inclusion of a method or research case study in the framework as an illustration of a useful technique does not imply panel acceptance of the chemical-specific outcome. Core panel members will serve for 2-3 years; members may be added to the standing panel to ensure expertise on specific topics.
- Science Panel members are selected from a diversity of affiliations and areas of expertise, particularly biology/toxicology, exposure assessment, epidemiology, risk assessment, and statistical/modeling. Core Panel members include:
  - James Bus, Exponent
  - Chris Chaisson, The Lifeline Group
  - Michael Dourson, TERA
  - Annie Jarabek, US EPA
  - Judy LaKind, LaKind Associates LLC (excused)
  - Bette Meek, University of Ottawa
  - Greg Paoli, Risk Sciences International

For workshop XI four *ad hoc* panelists have been selected. These scientists are:

- Carol Burns, Burns Epidemiology Consulting
- Krystin Carlson, US CDC
- Peg Coleman, Coleman Scientific Consulting
- Cissy Li, US FDA



Final Workshop XI Agenda & Purpose

To advance the recommendations in the NAS (2009) report concerning issue identification (problem formulation) and all aspects of risk assessment and management, through selection of illustrative research case studies for further development

Day 1: Tuesday, February 18th

RAAC Chair 1: **Pamela Williams,** E Risk Sciences

Welcome (1:00 to 1:15)

- **TBA,** NIOSH
- **Neeraja Erraguntla,** Member of the Risk Assessment Advisory Committee
- **TBA,** Member of the Science Panel

Keynote Talk (1:15 to 1:45)

- **Frank J. Hearl, Chief of Staff**
  National Institute for Occupational Safety and Health

Research Case Study 1: Occupational Exposure Banding 2.0: Characterizing Risks for Chemicals with Limited Data (1:45 to 3:00)

- **Christine Whittaker,** NIOSH
- Discussion by the Science Panel
- Comments from Observers

Afternoon Break (3:00 to 3:30)

Research Case Study 2: Use and Application of Real-Time Exposure Monitoring (3:30 to 5:00)

- **John Snawder,** NIOSH
- Discussion by the Science Panel
- Comments from Observers

Social TBA-open to all attendees
(dinner portion hors d'oeuvres, 5:30 to 8:00)

<<< >>>

Day 2: Wednesday, February 19th

RAAC Chair 2: **Pamela Williams,** E Risk Sciences and **Sabine Lange,** TCEQ

Research Case Study 3: Applying Hypothesis-Testing Methods to Help Inform Causality Conclusions from Epidemiology Studies (8:30 to 10:00)

- **Sabine Lange,** TCEQ

- Discussion by the Science Panel

Morning Break (10:00 to 10:30)

Research Case Study 3: continued (10:30 to 11:30)

- Discussion by the Science Panel
- Comments from Observers
- Chair's Summary

Ongoing Activities (11:30-noon)

- Data-derived Extrapolation Factors for Developmental Toxicity: A Preliminary Research Case Study with Perfluorooctanoate (PFOA)
  - **Bernard Gadagbui,** TERA

Lunch (noon to 1:00)

Ongoing Activities: continued (1:00-3:00)

- A Repair for Non-Cancer Assessment: Introducing the Truly Adverse Dose (the TAD)
  - **Lawrence Tannenbaum,** U.S. Army Public Health Center

- Proposal on the Epidemiology Peer Review Council
  - **John Doherty,** EPA (retired)

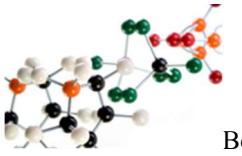
- Novel and Integrated approaches to modeling aggregate exposure to chemicals across different conditions of use and routes of exposure
  - **Clare Thorp,** Crème Global

- Bridging the Epidemiology Risk Assessment Gap: An NO_2 Case Study of the Matrix
  - **Carol Burns,** Burns Epidemiology Consulting

Afternoon Break (3:00 to 3:30)

Research Case Study 4: Understanding Weight-of-Evidence of Ototoxicity from Co-Exposures to Noise and Chemicals in the Workplace (3:30 to 5:30)

- **Neeraja Erraguntla,** ACC
- Discussion by the Science Panel
- Comments from Observers
- Chair's Summary



<<< >>>

Day 3: Thursday, February 20th

RAAC Chair 3: **Neeraja Erraguntla,** ACC

Research Case Study 5: Risk/Benefit Methods for Carcinogenicity/Sterilization with Ethylene Oxide as an Example (8:30 to 10:00)

- **Lucy Fraiser,** Lucy Fraiser Toxicology Consulting LLC
- Discussion by the Science Panel
- Comments from Observers
- Chair's Summary

Morning Break (10:00 to 10:30)

Research Case Study 6: Risk Assessment Methods of Flavoring in E-Vapor Products (10:30 to noon)

- **Julia Hoeng**, Philip Morris International and **Donna Smith**, Altria
- Discussion by the Science Panel

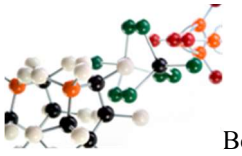
Lunch (noon to 1:00)

Research Case Study 6: continued: (1:00 to 3:00)

- Discussion by the Science Panel
- Comments from Observers
- Chair's Summary

Summary of the Workshop (3:00)

- **RAAC /Science Panel**

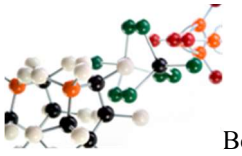


Biographical Sketches of Workshop Co-Chairs, Speakers, Presenters, and Science Panelists

Dr. Carol Burns is president of Burns Epidemiology Consulting with expertise in occupational and environmental epidemiology. Her research has addressed worker health, particularly in the context of evaluating and reducing exposure. Carol's primary area of focus is improving the use of epidemiology for use in risk assessment. She is passionate about communicating the role of epidemiology in decision making to colleagues, community members and government regulators. She holds a doctorate degree in epidemiology from the University of Michigan and a master's degree from Tulane University School of Public Health and Tropical Medicine. Active in the epidemiology community, she has held leadership roles in the American College of Epidemiology for many years. She serves on the editorial board of the journal Advances of Public Health.

Dr. James S. Bus is a Senior Managing Scientist in the Health Sciences Group of Exponent, Inc. (May 2013-present). Dr. Bus retired from The Dow Chemical Company as Director of External Technology and Fellow in the Toxicology and Environmental Research and Consulting unit (1989-2013). Prior to Dow, he was Associate Director of Toxicology and Director of Drug Metabolism at The Upjohn Company (1986-1989); Senior Scientist at the Chemical Industry Institute of Toxicology (CIIT, 1977-1986); and Assistant Professor of Toxicology, University of Cincinnati (1975-1977). Dr. Bus has been an advisor to a variety of institutions including ILSI, ILSI-HESI, The Hamner Institutes (formerly CIIT), American Chemistry Council Long-Research Initiative, and on advisory boards of the EPA (BOSC and Chartered SAB), FDA (NCTR), the National Toxicology Program, the National Academy of Sciences (BEST), and BELLE. He has served as President of the Society of Toxicology, The American Board of Toxicology, and the Academy of Toxicological Sciences, and in editorial roles including *Toxicology and Applied Pharmacology*, *Environmental Health Perspectives*, *Regulatory Toxicology and Pharmacology* and *Current Opinions in Toxicology*. Dr. Bus has received the Society of Toxicology Achievement (1987) and Founders (2010) awards, the Toxicology Forum George Scott Award (2013), Rutgers University Robert A. Scala Award (1999), the Michigan State University K.E. Moore Outstanding Alumnus Award, the International Society of Regulatory Toxicology and Pharmacology International Achievement Award (2015), and the International Dose-Response Society Outstanding Leadership Award (2018). He received a B.S. in Medicinal Chemistry from the University of Michigan (1971) and PhD in pharmacology from Michigan State University (1975), and currently is an Adjunct Professor in the Dept. Pharmacology and Toxicology at that institution. He has authored/co-authored over 150 publications, books, and scientific reviews. His primary research interests include modes of toxic action of industrial chemicals and pesticides including the role of non-linear toxicokinetics as a key consideration for improving the human relevance of *in vitro* and *in vivo* toxicity test findings.

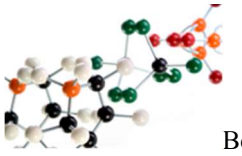
Dr. Krystin Carlson is a risk assessor at NIOSH with expertise in lead. She completed a PhD in Toxicology at the University of Michigan. Dr. Carlson's research focused on hearing loss in mice and humans (workers and newborns). Her research included investigating toxicants



mixtures. Projects included exposures to chemical (lead) with simultaneous exposure to a physical hazard (noise); multiple metal toxicants in combination; as well as essential metal interactions with toxicant metals. Dr. Carlson has been a professor at Oakland University in Michigan teaching graduate and undergraduate students. Adding to her contributions to the scientific community, Dr. Carlson has been an active member of the Society of Toxicology (SOT) since 2011 where she has served as a leader for graduate students as well as a representative for the SOT – Mixtures Specialty Section. As part of the SOT Graduate Student Leadership Committee, she served as both the Secretary and the Professional Development Chair. During these leadership positions she successfully organized a three-part seminar series on diversity and inclusivity in science. Dr. Carlson also won the SOT – Mixtures Specialty Section Best Abstract Award for her mouse model on investigations of ototoxicity due to lead, cadmium, and noise. Dr. Carlson completed her undergraduate degree in Environmental Health Sciences at Purdue where she participated in research exploring occupational mortality, hospital noise, ergonomics, and food science.

Dr. Christine Chaisson is a Director in The LifeLine Group™ and a senior member of the LifeLine Group's management team. She is one of key architects of the new generation of exposure assessment models addressing aggregate and cumulative risk concepts, called LifeLine™. Dr. Chaisson earned a doctorate in cellular biochemistry/biology from George Washington University (1982). She began her career in risk assessment in the US Environmental Protection Agency in the Office of Pesticides and Toxic Substances. At EPA, Dr. Chaisson designed and created the first probabilistic dietary exposure assessment model. She was also the liaison to international regulatory agencies such as AID and WHO. In 1985, Dr. Chaisson co-founded Technical Assessment Systems (TAS), which became the premier exposure/risk assessment consulting firm internationally. Through TAS, she introduced concepts such as population subgroup specificity, better definition of residues in forms of foods and sources of drinking water, use of human activity patterns and actual chemical usage patterns for more accuracy and relevance in risk assessment models. Through these experiences, Dr. Chaisson became well versed in the expectations of regulators in the US, UK, Canada, Germany and European Union.

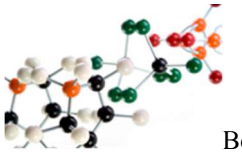
Dr. Chaisson has been a Councilor in the International Society of Exposure Assessment, a member of Society of Risk Assessment and President of its DC chapter, the Toxicology Forum, the United Agribusiness League and the Institute of Food Technologists. She also served on the National Council for Arts and Sciences of the George Washington University and the Dean's Advisory Board for the GWU Graduate School of Political Management. Dr. Chaisson serves as a member of the External Advisory Board of the Center for Indigenous Environmental Health Research at the Zuckerman College of Public Health / University of Arizona. She is an advisor to Food Quality magazine. She has published extensively in the fields of exposure and risk assessment. In 2011 Dr. Chaisson was the invited Co-Chair of the Milan ISES/SETAC special conference on exposure science challenges presented by global legislative initiatives on consumer products and chemicals in trade. In 2014 Dr. Chaisson led a multi-presentation ISES session and panel presenting the Community Based Research in post-Sandy Brooklyn to characterize clean-up workers' exposure to industrial chemicals displaced by the storm. A follow-up symposium on that and related work was presented at the 2018 ISES-ISEE joint meetings in Ottawa, Canada.



Ms. Peg Coleman, MS[2], is a consultant in microbial risk assessment who provides novel perspective to the Science Panel for this workshop on the key roles of the gut microbiota in evolving paradigms about chemical risk assessment. She currently serves as secretary/treasurer of the Dose-Response Specialty Group of the Society for Risk Analysis. Her career as a microbial risk assessor began with the US federal government (USDA), she consulted through Syracuse Research Corporation and ICF before launching her own firm in 2010. She has collaborated with gifted mathematical statisticians throughout her long career and is an author on 25 peer-reviewed manuscripts, most collaborations on microbial risk assessment (including one submitted this year and another in preparation) and one each in chemical risk assessment and immunology. Ms. Coleman's current interests include benefit-risk assessment, incorporating whole genome sequencing and metadata into risk assessment, and the roles of natural and engineered microbiota of foods and humans relevant to health and disease for humans and the environment**.**

Dr. John Doherty, a specialist in Neurotoxicology**,** earned his BS from the University of New Hampshire (Biochemistry), his MS- from George Washington University (Biochemistry) and his Ph.D. from the University of Wisconsin (Entomology/Biochemistry – mode of action of DDT). He was a Postdoctoral Fellow at- Yale University Department of Pharmacology (Natural occurrence and metabolism of gamma-hydroxybutyric acid). Throughout his career, Dr. Doherty has worked as a Chemist for the Food and Drug Administration, and Assistant Professor at the University of Chicago Department of Psychiatry (Neuropharmacology of phencyclidine) and as a Toxicologist in the– Health Effects Division, Office of Pesticide Programs, USEPA. He is a Diplomat of the American Board of Toxicology (1982-2017) and was a Visiting Scientist at National Institutes of Health (inhibition of ATPases by pesticides, lead (Pb) and possible interaction with omega-3-fatty acid) as well as a Visiting Scientist at Kyoto University (Japan) (non-sodium dependent release of transmitters by pyrethroids in synaptic preparations and inhibition of Ca++),, the University of Colorado and the University of North Carolina (dopaminergic neuropharmacology of triadimiform). He is currently retired (since 2012).

Dr. Michael Dourson has a PhD in toxicology from the University of Cincinnati, College of Medicine, and is a board-certified toxicologist (i.e., DABT) serving as the Director of Science at the 501c3 nonprofit organization Toxicology Excellence for Risk Assessment (TERA). Prior to this, he was Senior Advisor in the Office of the Administrator at the US EPA. Before this, he was a Professor in the Risk Science Center at the University of Cincinnati, College of Medicine and also worked at TERA and US EPA. He has been awarded the Arnold J. Lehman award from the Society of Toxicology, the International Achievement Award by the International Society of Regulatory Toxicology and Pharmacology, and 4 bronze medals from the U.S. Environmental Protection Agency. He has been elected as a Fellow of the Academy of Toxicological Sciences (i.e., FATS) and as a Fellow for the Society for Risk Analysis (i.e., FSRA). He has co-published more than 150 papers on risk assessment methods or chemical-specific analyses, and co-authored well over 100 government risk assessment documents, many of them risk assessment guidance texts. He has made over 150 invited presentations to a variety of organizations, and has chaired over 150 sessions at scientific meetings and independent peer reviews. He has been elected to multiple officer positions in the American Board of Toxicology (including its President), the Society of Toxicology (including the presidency of 3 specialty sections), the Society for Risk Analysis (including its Secretary), and is currently the President of the Toxicology Education



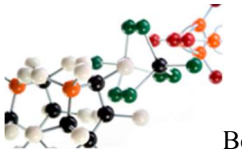
Foundation, a nonprofit organization with a vision to help our public understand the essentials of toxicology. In addition to numerous appointments on government panels, such as EPA's Science Advisory Board, he is a current member on the editorial board of Regulatory Toxicology and Pharmacology and Human and Experimental Toxicology.

Dr. Neeraja K. Erraguntla (Neera) is a Director, at the Chemical Products and Technology division at the American Chemistry Council (ACC). Dr. Erraguntla is responsible for managing and directing ACC's 1,3-Butadiene TSCA Risk Evaluation Consortium and the Center for Advancing Risk Assessment and Science policy under ACC's Center for Chemical Safety. In addition, she also manages four other industrial chemical groups that endeavor for the development and application of up-to-date, scientifically sound methods for conducting chemical assessments. Dr. Erraguntla directs complex projects involving systematic reviews, mode-of-action, exposure characterization, and endocrine disruption.

Prior to ACC, Dr. Erraguntla was a senior regulatory toxicologist at the Texas Commission on Environmental Quality (TCEQ) from 2005 to 2015. At TCEQ, she was a team lead to review available tools for conducting systematic reviews and evidence integration and to develop a position paper on how TCEQ conducts systematic reviews and evidence integration. Neera also determined inhalation toxicity factors of arsenic compounds and hexavalent chromium compounds, and used threshold of concern to determine acute toxicity for chemicals with limited toxicity information. Neera played a major role in understanding and addressing community concerns about increased asthma rates in children and adults and prepared several science-based regulatory evaluations.

Dr. Erraguntla is a diplomate of American Board of Toxicology (DABT) and has a Ph.D. from Louisiana State University. She volunteers with SOT Risk Assessment Specialty Section and has also volunteered and served on the committee for SOT Exposure Specialty Section. Dr. Erraguntla was nominated as a Council Member for the International Society of Regulatory Toxicology & Pharmacology (ISRTP). In 2016, she served as a reviewer for the Government's Accountability Office and was a peer reviewer of the National Academies report, Acute Exposure Guideline Levels for Selected Airborne Chemicals, Volume 20, from the Board on Environmental Studies and Toxicology. Previously, Dr. Erraguntla also served as a Science Advisory Board (SAB) for US EPA's Environmental Justice Technical Guidance Panel and has been on the National Academy of Sciences Acute Exposure Guidelines Committee. Previously, she served as an Adjunct Assistant Professor at Texas A&M School of Public Health.

Dr. Lucy Fraiser is the Principal Toxicologist and Owner of Lucy Fraiser Toxicology Consulting LLC. She has over 30 years of experience and a proven track record in the areas of exposure and risk assessment, health effects and toxicology evaluations, development of quantitative toxicity criteria, development of risk-based air and water quality guidelines and soil cleanup criteria, and risk communication. While Dr. Fraiser works with all environmental media, she specializes in air quality health evaluations. Dr. Fraiser is well known for being a thorough and skilled scientist with good writing skills and an excellent communicator. She is praised by her repeat clients for her ability to conduct in-depth analyses of the scientific literature and attention to detail, as well as her ability to clearly and confidently communicate complex



toxicological and health risk concepts in a manner that is understandable and persuasive. Dr. Fraiser has examined the scientific foundation on which exposure assumptions and toxicity criteria are based on many occasions on behalf of private and public-sector clients and trade organizations. Her leading work on these issues has resulted in corrections to regulatory guidance and risk-based criteria on several occasions. Dr. Fraiser is actively involved in ongoing discussions/debates about the potency of ethylene oxide as a carcinogen. As the only toxicologist on Task Forces for two municipalities with ethylene oxide medical supply sterilization facilities, she provided specialized knowledge and expertise on the risk potentially associated with air monitoring results and helped develop strategies for the assessment of actions taken by the Agency for Toxic Substances and Disease Registry (ATSDR) and the Environmental Protection Agency (EPA). She testified before the Illinois House and Senate on ethylene oxide risk and cancer potency on behalf of the Advanced Medical Technology Association (AdvaMed) and she will be participating in a Congressional Briefing on ethylene oxide on February 19, 2020.

Dr. Bernard Gadagbui joined Toxicology Excellence for Risk Assessment (TERA) since 2004 and is currently is a Senior Toxicologist at TERA, with extensive experience in toxicology and human health risk assessment. Dr. Gadagbui received a BSc in Biochemistry with Chemistry from the University of Ghana, Legon, Ghana, and MSc in Biochemistry and PhD in Environmental Health from the University of Bergen, Norway. He has sound understanding of toxicology/risk assessment principles/practices, scientific basis for toxicity testing guidelines and application of science-based risk assessment methodologies. His extensive evaluation of clinical and non-clinical data and use of read across approaches has resulted in derivation of numerous high quality toxicologically-based risk values including reference doses/concentrations, occupational exposure limits, acceptable daily exposures, and permitted daily exposures for data-rich and data-poor chemicals, including industrial chemicals, manufacturing reagents, pesticides, cosmetic and personal care ingredients and products, botanicals and botanical preparations, petroleum hydrocarbons, and active and inactive pharmaceutical ingredients. Dr. Gadagbui is certified as a Diplomate of the American Board of Toxicology (DABT) and is also a European Registered Toxicologist (ERT). He has held leadership positions in the Toxicologists of African Origin (TAO), a Specialty Interest Group of the Society of Toxicology (SOT), African Society of Toxicological Sciences (ASTS), Ohio Valley Chapter of SOT, Ohio Chapter of Society for Risk Analysis (SRA) and currently one of the three Advisors of the recently formed African Chapter of SRA (SRA-Africa).

Mr. Frank Hearl, P.E joined the National Institute for Occupational Safety and Health (NIOSH) in 1974 and is presently the NIOSH Chief of Staff. He assists the NIOSH Director with provide broad oversight and management of the Institute's research and service activities. He is the Institute's point-of-contact to coordinate with other federal agencies such as the Occupational Safety and Health Administration (OSHA), the Mine Safety and Health Administration (MSHA), and the Environmental Protection Agency (EPA), promoting the translation of NIOSH research and risk assessments to practice for science-based decision making.

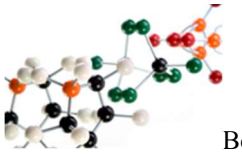
His past experience includes conducting field industrial hygiene and epidemiology studies related to coal workers pneumoconiosis, silicosis, lung cancer, and various lung disease-producing agents. He was the NIOSH lead on an international project with the Tongji Medical

University in China and the U.S. National Cancer Institute to study silica, silicosis and lung cancer. In his 30-years of service as a U.S. Public Health Service Officer, he received 17 awards including the Meritorious Service Medal. He has been twice awarded the Alice Hamilton Award for excellence in Occupational Safety and Health, and the Bullard-Sherwood Award for advancing "Research to Practice."

Mr. Hearl received his bachelor's degree in chemical engineering from Purdue University in 1974, and earned a master's degree in chemical engineering from the Massachusetts Institute of Technology (MIT) in 1980. He is licensed Professional Engineer in Maryland and West Virginia.

Dr. Julia Hoeng is Director of Systems Toxicology at Philip Morris International Research & Development where she leads the Systems Toxicology Biology Program, covering a portfolio of projects from *in vitro*, *in vivo* and *in silico* research for product testing. Dr. Hoeng has established a multidisciplinary team leveraging most recent developments in molecular measurements and high throughput screening methods, in-vitro aerosol testing, computational toxicology and bioinformatics. Dr. Hoeng holds several patents related to protein biochemistry, computational biology, biotechnology and has published over 100 peer reviewed scientific articles, book chapters and invited strategic reviews in the field of Systems Toxicology, Inhalation Toxicology, Crowd Sourcing, Molecular and Computational Biology. She holds a PhD and Post-doc from Cambridge University and a MS in Bioinformatics from the Georgia Institute of Technology, Atlanta, Georgia, USA.

Ms. Annie M. Jarabek currently serves as a Senior Science Advisor in the immediate office of the Center for Public Health and Environmental Assessment (CPHEA) at its Health and Environmental Effects Assessment Division (HEEAD) in the Research Triangle Park, within the U.S. Environmental Protection Agency's Office of Research and Development (ORD), following recent service as the Deputy Director of the Human Health Risk Assessment (HHRA) national research program in ORD. Ms. Jarabek has significant experience and training in inhalation toxicology in both laboratory and clinical environments, dosimetry modeling, risk assessment, and decision analysis. She was principal author of the Agency's *Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry*. Ms. Jarabek has worked on risk assessments, dosimetry models or analysis methods across all media and routes of exposure. She was the lead for the Agency's risk assessment of ingested perchlorate and some of her other work addressed several priority interdisciplinary Agency assessments including: inhaled particulate matter, vinyl acetate, manganese, and Libby amphibole asbestos. Her current research efforts focus on multi-scale dosimetry modeling, including approaches for *in vitro* to *in vivo* extrapolation (IVIVE) of inhalation exposures to advance the application of emerging methods for translation and evidence integration across various experimental platforms. A manuscript on her collaborative IVIVE work received an honorable mention as the best 2018 paper from the Biological Modeling Specialty Section (BMSS) at the 2019 annual Society of Toxicology (SOT) meeting. Ms. Jarabek has received three awards for best manuscript in risk assessment application from the Risk Assessment Specialty Section (RASS) of the SOT, along with several best abstract presentation awards. She has also received a Lifetime Achievement Award from the University of Massachusetts, the Risk Practitioner of the Year award from the Society of Risk Analysis (SRA), the Superfund National Notable



Achievement Award, and several award medals (gold, silver and bronze) and technical or special service awards from the Agency. She will be awarded the Lehman award for risk assessment at the 2020 SOT meeting in Baltimore.

Dr. Sabine Lange is the section manager for the Toxicology Division at the Texas Commission on Environmental Quality (TCEQ). Dr. Lange's responsibilities include overseeing health effects risk assessments of air permit applications, ambient air monitoring projects, and hazardous waste sites; overseeing the development of chemical toxicity factors; and conducting and overseeing systematic reviews and independent analyses of risk assessments. Dr. Lange serves as a technical resource for the State and citizens of Texas for human health and environmental risk assessment, especially related to air and water quality. Dr. Lange's research interests include the toxicology of criteria air pollutants, and risk assessment methods used for derivation of toxicity factors. Dr. Lange received a bachelor's degree from the University of Western Ontario in Canada, and completed a Ph.D. and post-doctoral training in biochemistry and molecular carcinogenesis at the University of Texas at Houston and MD Anderson Cancer Center. Dr. Lange is a Diplomate of the American Board of Toxicology.

Dr. Cissy Li is a Toxicologist at the US Food and Drug Administration (FDA) Center for Tobacco Products (CTP). Her work at CTP involves tobacco product application review and toxicological research, and her research focus is e-cigarettes. She previously worked as an ORISE Fellow at the FDA Center for Food Safety and Applied Nutrition, where she evaluated chemical contaminants in food. Dr. Li completed her doctoral training in toxicology at the Department of Environmental Health Sciences at the Johns Hopkins Bloomberg School of Public Health, and obtained a certificate in Risk Sciences and Public Policy from the same institution. Her experiences include risk assessment and *in vitro* and *in vivo* models in academic and regulatory settings.

Dr. Bette Meek is the Associate Director of Chemical Risk Assessment at the McLaughlin Centre for Risk Science, Faculty of Medicine, University of Ottawa. Previously, she contributed to and managed several chemical risk assessment programs within Health Canada. With colleagues internationally, she has contributed to or led initiatives in developing methodology in chemical risk assessment, including mode of action, chemical specific adjustment factors, physiologically-based pharmacokinetic modeling, combined exposures and predictive modeling. These initiatives have involved collaborations with a range of international organizations and national Agencies, including the World Health Organization International Programme on Chemical Safety, the Organization for Economic Cooperation and Development, the U.S. Environmental Protection Agency, the European Joint Research Centre and the Agency for Food, Environmental and Occupational Health and Safety of France (ANSES). She has authored approximately 200 publications in this area and received several awards for contribution in this domain.

Dr. Greg Paoli's career has spanned a wide spectrum of public risk management domains. This has included the safety of food, drinking water, air quality, consumer products, drugs, medical devices and the blood supply, engineered devices, transportation of dangerous goods, museum collections, emergency management for natural and man-made disasters, and climate change

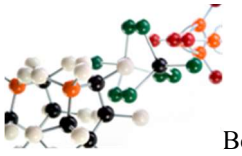
impacts on infrastructure. Due to the diversity of this experience, Greg was commissioned by the University of Pennsylvania Law School to prepare a discussion paper on "The Analytical Capabilities of a Best-in-Class Regulator" as part of its international Best-in-Class Regulator Project.

Dr. Paoli has served on a number of expert committees devoted to the risk sciences. He was a member of the U.S. National Academy of Sciences committee that issued the 2014 report, *A Framework to Guide the Selection of Chemical Alternatives*, and the 2009 report, *Science and Decisions: Advancing Risk Assessment*. He was invited to serve as a member of an expert peer review panel for the US EPA's Framework for Human Health Risk Assessment to Inform Decision Making. He has served on numerous expert committees convened by the World Health Organization and the Food and Agriculture Organization of the United Nations. He recently served a three-year term on the Scientific Advisory Committee for Health Canada's Chemical Management Plan.

Dr. Paoli completed a term as Councilor of the Society for Risk Analysis (SRA) and served two terms as a member of the Editorial Board of the journal Risk Analysis. In 2011, he was awarded the Distinguished Lectureship Award by the Society for Risk Analysis and the scientific society, *Sigma Xi*.

Dr. Davide Sciuscio earned a master's degree in Pharmaceutical Chemistry (2005) and later a PhD in Pharmacology and Toxicology (2008) from University of Bologna, Italy. Currently he holds the position of Manager of the Pre-clinical and Toxicological Evaluation at Philip Morris International (PMI). In his role as a toxicologist, Dr. Sciuscio is in charge of the risk assessment of chemicals and ingredients as well as for the toxicological and pre-clinical evaluation of Electronic Nicotine Delivery Systems (ENDS) and Heated Tobacco Products (HTPs). In particular, by performing both *in vitro* and *in vivo* studies, Dr Sciuscio and his team provide scientific support to product development, clinical studies, regulatory submissions, and product market launches. Prior to joining Philip Morris international in 2012, Dr. Sciuscio held a post-doctoral position in the laboratory of brain tumor biology and genetics at the University of Lausanne, Switzerland. His research was focus on the application of OMICs technologies toward a better understanding of brain tumors biology in order to provide new diagnostic/predictive tools and ultimately new molecular targets for future treatments. Dr. Sciuscio is a member of the Italian Register of Toxicologists and a European Registered Toxicologist (ERT since 2014). He is a member of Swiss and Italian Society of Toxicology (SST and SITOX) and he is author of 15 scientific publications and has presented at many international meetings over the last decade.

Dr. Donna C. Smith, Ph.D., DABT serves as Associate Fellow, Preclinical within Altria Client Services' Regulatory Sciences organization. She leads a group of toxicologists responsible for the toxicological review of all ingredients, prototypes and products across Altria's tobacco product portfolio. Her group spans risk assessment activities to biological testing and includes significant subject matter expertise included in substantive pre-market product applications to the Center for Tobacco Products within FDA.



Dr. Smith has held a variety of positions within the Altria family of companies with increasing levels of accountability within the preclinical organization since 1999. Most recently she served as Sr. Principal Scientist and Principal Scientist where she led the development and utilization of toxicological risk assessments to support product development and facilitated regulatory engagement and compliance.

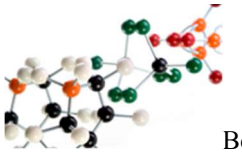
Dr. John E. Snawder received his Ph.D. degree from Mississippi State University in 1990 and conducted post-doctoral work in the National Center for Toxicological Research. He has been a Diplomate of the American Board of Toxicology since 1995 and served on their Board of Directors from 2008 to 2012. From 1992-present he has worked at NIOSH as a research toxicologist. Dr. Snawder is the co-Director of the NIOSH Center for Direct Reading and Sensor Technology and is the Leader of the Biological Monitoring Research Team. Dr. Snawder and his team coordinate biological monitoring for many NIOSH studies and are presently involved in exposure assessment and biomonitoring studies for upstream oil and gas workers, fuel handlers, nanotechnology workers and first responders.

Dr. Lawrence Tannenbaum is a biologist and certified senior ecologist who has worked for the past 25 years as a senior health risk assessor for the Army Public Health Center, located at Aberdeen Proving Ground, Maryland. Prior to joining the Army, he was a human health and ecological risk assessor for the U.S. EPA in that agency's Region 2. He has published some 50 times in the peer-reviewed literature on varied risk assessment topics, and has published two solo-authored books on ecological risk assessment, the first of which -- "Alternative Ecological Risk Assessment" -- is housed in close to 600 university libraries in the U.S. and abroad. His area of high interest is in demonstrating the absence of ecological impacts at contaminated sites, and in accounting for that phenomenon. His formerly patented Rodent Sperm Analysis method recently attained ASTM International certification, and stands as the only field-based ASTM method in the ecological arena and for the Society. For the past 15 years, Dr. Tannenbaum has been a senior editor for the SETAC journal, Integrated Environmental Assessment and Management, and he recently joined the journal, Drug and Chemical Toxicology as a Section/Associate editor.

Dr. Clare Thorp is Senior Advisor to Creme Global* for their North American market. She is also CSO and Co-Founder of Oaklare Management Corporation, which provides consultancy services to life-science companies.

Creme Global is an advanced software, data analytics and scientific modelling company that was established over 15 years ago and is headquartered in Dublin, Ireland. Creme Global works with industry, government and academia, providing insights into large and complex data sets in support of better decision making. They are recognized for their leadership and advancement of exposure science, and their ability to curate and manage highly sensitive data sets across multiple industry partners.

Dr. Thorp's diverse expertise and experience is built on a career in the United States and European Union, where she has worked as an academic, in government and with various industry sectors. Her expertise is underpinned by a degree in agricultural science, a Ph.D. in animal

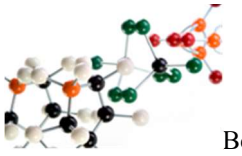


science, a postgraduate certificate in European law, and a master's degree in economics and policy. She can also drive a tractor and knows how to farm.

Dr. Thorp first came to the United States in 2005, as a senior diplomat at the Embassy of Ireland, Washington D.C. where she was tasked with establishing the new Food and Agriculture Office on behalf of the Irish government. Since 2010, Dr. Thorp has held executive leadership positions in trade associations and non-profits including CropLife America, the Biotechnology Innovation Organization and the International Life Sciences Institute of North America. She brings a unique and expert understanding of how science, policy and regulation intersect, coupled with an ability to understand and address issues from a variety of perspectives, which enables cross cultural problem solving and collaborative issue management.

Dr. Patrick Vanscheeuwijck is Director Pre-clinical Toxicology at Philip Morris International, Life Sciences, Neuchatel, Switzerland, responsible for the *in vitro* and *in vivo* assessment of Reduced Risk Products (RRPs), with leadership over research and assessment groups in Neuchatel and Singapore. The focus of his 25 year career at PMI has been on the development of approaches for the assessment of hazard associated with cigarette smoke and aerosols from RRPs, *in vitro* research and assessment, inhalation toxicology and animal models of disease; with more than 55 peer-reviewed publications. He started his career at PMI as research scientist, then headed the Bioresearch group and became the General Manager of the *in vivo* testing laboratories, Philip Morris Research Laboratories in Leuven, Belgium (75 headcount). In 2010, he became the General Manager of the Philip Morris Research Laboratories in Cologne, Germany (210 headcount) with focus of analytical chemical, bio-analytical, and *in vitro* assessment of cigarette smoke and aerosols from potential modified risk products . As from 2013 he moved to the R&D Research and Development organization in Neuchatel, Switzerland where he was in the position of Head of Experimental Biology and subsequently as Director Pre-clinical Toxicology. He obtained his Masters degree at the University of Gent, Belgium, in biochemistry and a Ph.D. in biochemical pharmacology (University of Gent, Belgium), performed postdoctoral fellowships in molecular pharmacology and receptor biology at the University of Arizona, Tucson, U.S. and in molecular biology at the developmental biology group at the University of Leuven, Leuven, Belgium.

Dr. Christine Whittaker received her B.S. in Biology from Rensselaer Polytechnic Institute and her Ph.D. in Environmental Toxicology from the University of California Irvine. From 1990 to 1997, Dr. Whittaker worked in the Directorate of Health Standards in the Occupational Safety and Health Administration (OSHA) in Washington DC, where she conducted occupational risk assessment to support chemical regulations. In 1997, Dr. Whittaker moved to the NIOSH Office of the Director in Washington, DC, where she served as a senior scientist. In 2004, she moved to Cincinnati, Ohio, as Chief of the Risk Evaluation Branch in the Division of Science Integration (DSI). Throughout her career, Dr. Whittaker's focus has been assessing chemical hazards to workers and determining how those hazards can be most effectively mitigated through science policy. In her time in DSI, Dr. Whittaker has been involved in the development of NIOSH science policy, including thinking around the utility and processes in systematic review, the NIOSH Chemical Carcinogen Policy, the NIOSH Practices in Occupational Risk Assessment,



the NIOSH Occupational Exposure Banding Process for Chemical Risk Management, as well as several documents on occupational exposure to various workplace chemicals.

Dr. Pamela Williams is a Principal at E Risk Sciences, LLP, an independent scientific consulting firm that provides sound analyses and tools to support risk-based decision-making related to human health and the environment. She is also a Clinical Assistant Professor in the Department of Environmental and Occupational Health at the Colorado School of Public Health as well as a Fellow with the non-profit organization Toxicology Excellence for Risk Assessment (TERA). Dr. Williams specializes in assessing human exposures and health risks in environmental, community, and occupational settings. Her particular areas of expertise include human health risk assessment, exposure science, exposure modeling, and uncertainty analysis. She has published over 100 papers, book chapters, and presentation abstracts on various risk-related topics. She has also taught graduate-level and continuing education courses related to exposure and risk assessment at the Colorado School of Public Health, Harvard School of Public Health, Society of Toxicology, and the American Industrial Hygiene Association (AIHA). She routinely serves as a technical peer-reviewer for a number of scientific journals, peer review panels, and government agencies. Dr. Williams is past President of the Society for Risk Analysis (SRA) and past Chair of AIHA's Risk Committee. She has received several awards for her contributions to the fields of risk analysis, exposure science, and industrial hygiene. These include the Chauncey Starr Distinguished Award granted by the Society for Risk Analysis for excellent contributions to the field of Risk Analysis, the Joan M. Daisey Outstanding Young Scientist Award granted by the International Society of Exposure Science for outstanding contribution to the science of human exposure analysis, and both a Leadership Award and Outstanding Individual Contributor Award granted by AIHA in recognition of leadership and outstanding contributions to AIHA. Dr. Williams has a B.A. in Sociology and Applied Social Research from San Diego State University, M.S. in Health and Social Behavior from Harvard University, and ScD in Environmental Health and Health Policy and Management from Harvard University. She is also a certified industrial hygienist (CIH).