UNITED STATES OF AMERICA

DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

+ + +

CENTER FOR TOBACCO PRODUCTS

+ + +

RISK ASSESSMENT OF TOBACCO PRODUCTS: A PUBLIC WORKSHOP

+ + +

November 16, 2016
8:30 a.m.

The Marriott Inn and Conference Center
University of Maryland University College (UMUC)
3501 University Boulevard East
Hyattsville, MD 20783

FDA:

CAROLYN DRESLER, M.D., M.P.A.
Associate Director for Medical and Health Sciences
Office of Science
Center for Tobacco Products

PUBLIC COMMENT SESSION

ALI A. ROSTAMI, Ph.D.
ALCS Research Development & Regulatory Affairs
Altria Client Services

BRUCE LEVINSON
Center for Regulatory Effectiveness

SESSION 4: RISK ASSESSMENT

P. ROBINAN GENTRY, Ph.D., DABT
Ramboll Environ, Inc.

VINCENT COGLIANO, Ph.D.
Director, IRIS Program
National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency

PETER SHIELDS, M.D.
Ohio State University

GLENN RICE, Sc.D.
National Center for Environmental Assessment
U.S. Environmental Protection Agency

STEPHEN ROBERTS, Ph.D.
Center for Environmental & Human Toxicology
University of Florida

SESSION 5: RISK ASSESSMENT: CANCER AND NONCANCER APPROACHES

STEPHEN ROBERTS, Ph.D.
Center for Environmental & Human Toxicology
University of Florida

GEORGE GRAY, Ph.D.
Center for Risk Science and Public Health
Department of Environmental and Occupational Health
Milken Institute School of Public Health
The George Washington University

ANNETTE SANTAMARIA, Ph.D., M.P.H., DABT
Director of Toxicology
Rimkus Consulting Group, Inc.

SESSION 6: RISK ASSESSMENT: MIXTURES

MOIZ MUMTAZ, Ph.D.
Division of Toxicology and Human Health Sciences
Agency for Toxic Substances and Disease Registry
Centers for Disease Control and Prevention

LYNNE HABER, Ph.D., DABT
Toxicology Excellence for Risk Assessment Center
University of Cincinnati

SESSION 7: RISK ASSESSMENT: UNCERTAINTY AND VARIABILITY

RICHARD C. HERTZBERG, Ph.D.
Biomathematics Consulting
Emory University

KENNETH M. PORTIER, Ph.D.
VP Statistics & Evaluation
American Cancer Society

DALE HATTIS, Ph.D.
Clark University

M E E T I N G

(8:32 a.m.)

DR. DRESLER: My name is Carolyn Dresler, and I'm the moderator for this Risk Assessment of Tobacco Products from the Center for Tobacco Products and Office of Science. And I know we're starting just a few minutes late. I was trying to hold on a little bit, but there's many people online, so it's not fair for them to wonder, you know, what the heck's going on in the room, so let's go ahead and get started.

Welcome back. We had an exciting and long day yesterday, but with a lot of good presentations and discussion, so I'm looking forward for that for today, too. We're going to start out first with the public comment session, and so we have speakers that will come up and present for us. So could we please have the first public speaker?

DR. ROSTAMI: Thank you.

Good morning. My name is Ali Rostami. I'm an Associate Principal Scientist at Altria Client Services. I'm going to talk about using computational modeling as a methodology for characterizing exposure to aerosols from electronic nicotine delivery systems.

One of the objectives of this workshop is to identify method development which may further strengthen knowledge

regarding tobacco products risk assessment. One specific question is what methods are available to characterize exposure to constituents associated with the use of different types of tobacco products?

One way to answer this question is to track the aerosol from its source to the endpoint of relevance to exposure and use computational modeling as a method to quantify the level of exposure. The end goal is to strengthen scientific knowledge base for product and exposure evaluations.

This diagram shows the flow path of aerosol from its point of formation to the endpoint relevant exposure. The aerosol is generated in the device, it is inhaled by the user, and part of the inhaled aerosol and constituents is absorbed and distributed in the body, and the remaining is exhaled into a confined space, which is relevant to the passive vaping. Unlike cigarettes, where the sidestream smoke contributes to the passive vaping, in e-cigarettes, only the exhaled aerosol is the contributor.

There are many variables along this path that impact the level of exposure. For example, device design variations, e-liquid composition, users' conditions, and indoor space conditions. Running multiple experimental studies to allow for the combination of these variables is impractical. As a

result, we are using, in parallel computational, we developed computational models to evaluate multiple scenarios.

There are four models that are listed here: The e-vapor characterization model, it calculates the amount, composition, and temperature of aerosol coming out of the device. And the aerosol deposition model estimates the rate of deposition of different constituents in the respiratory tract. The room air quality model calculates the concentrations of exhaled aerosol in the room. And the physiologically based PK model estimates the distribution of the constituents in the body.

These are not data-driven statistical models; they're physics-based models that include fluid, mass, and heat transfer along with thermodynamic and kinetic interactions.

Let's take a look at three of these models and sample results. The vapor characterization model takes in variables such as device design and value specification, the amount and composition of e-liquid, and the use conditions, like puff volume and duration, and it predicts the aerosol mass on puff-by-puff basis, as well as the heater and the aerosol temperature coming out of the device.

An example result from that model is the aerosol mass as a function of power input to the system. The line is produced by the computational model for Prototype 1 independently of the

data points. The data points here are the measured value of aerosol delivered by two different prototype devices. The agreement is remarkably good for Prototype 1 for which the model was developed, and it is very good for another device, which is a larger device but similar design configuration.

Another important performance parameter of a device is the heater temperature. The heater temperature affects the amount, composition, and temperature of the aerosol coming out of the device, which all are relevant to the exposure.

The next model is the aerosol deposition model. This model takes in variables like respiratory tract geometry, the device/aerosol delivery, the use condition, and calculates the flow conditions, temperature, humidity in the respiratory tract, as well as the deposition of individual constituents in the respiratory tract. More information on this model are available in our recent two publications listed down there.

And this model, as well as the other models, this one accounts for the complex droplet vapor physical and thermodynamic interactions presenting this in the e-vapor aerosol. An example result from this model is the PG, glycerin, nicotine deposition in the upper airways down to Generation 3, and you can see that the rate of deposition of PG and nicotine is higher than glycerin because of their

volatilities, higher volatilities compared to glycerin.

The last model is the air quality or passive vaping model that takes in variables such as amount of exhaled aerosol, the room size, and the ventilation rate, and calculates the room concentration of each of the constituents exhaled into the room as well as the spatial and temporal distribution in the room. More information on this model is also available in the recent publication listed here.

One result from this model is the nicotine vapor distribution, exhaled nicotine vapor distribution in a typical office space. This office space has a volume of 24 m^3, and the air ventilation rate is 0.35 air change per hour. In the first 30 minutes, 10 puffs are taken 3 minutes apart, and this movie shows the spatial distribution of the exhaled nicotine in the room and how it is carried out by the ventilation. Each plume is one exhaled puff. Also, you can see the average values of nicotine, PG, and glycerin over this 60 minutes. The numbers are relatively low for this situation.

Of course, an exposure time and duration is an important parameter. This graph shows for the same setting the PG, glycerin, and nicotine concentration, average concentration of the room over 60 minutes during the first 30 minutes when the e-cigarette is being used. It is rising; after that, all of

them start to decrease due to the ventilation and dilution with the room air.

In summary, in order to characterize exposure to ENDS aerosol constituents, it is important to check the aerosol from its source to the endpoints. Computational modeling can be used as a methodology for this characterization. We developed several models that track ENDS aerosol from the point of formation to the endpoint relevant exposure. These models are listed here: e-vapor characterization model, aerosol deposition model, room air quality model, and physiologically based PK model.

Thank you.

(Applause.)

DR. DRESLER: Okay. And then our second speaker, please.

MR. LEVINSON: Hi. I'd like to thank FDA for making it possible for me to be here today. Transparency in Washington has -- is somewhere between an obligatory buzzword and an empty cliché, but not at FDA. The Agency has done a tremendous amount of outreach of involving and making an open participatory process not just at this workshop, but at previous ones and throughout the entire tobacco risk assessment process.

I would like to thank the Agency but also suggest that the

Agency write this up, write up the entire risk assessment process as a case study. FDA isn't the only agency dealing with this sort of tremendously, devilishly complex long-term process. I would also like to thank FDA for making it possible for me to talk about both NASCAR and the Hell's Angels. Now, NASCAR, as you may know, started with bootlegging. It was a way that bootleggers souped up their cars to go beat out the revenues law enforcement authorities. Motorcycle gangs, at least some of them, also have a similar history, but with methamphetamine. The street term is crank, came from motorcycle gangs smuggling in the crankcase of their motorcycles.

So perhaps it's not surprising that a NASCAR driver, along with the Quebec chapter of the Hell's Angels, teamed up on a major international narcotics ring, and while that ring did traffic in narcotics, mostly it was centered on trafficking in tobacco. More than 50 million kilos of North American-grown tobacco went up into Canada, where it came back as cigarettes.

We're talking about two, over two billion cigarettes, and that's a lot of cigarettes. Now, we know that illicit tobacco has a different risk profile than legal products. There was a recent study, Kim Quayle and her colleagues in the September 2016 *Journal of Forensic Science International,* on how to tell

licit from illicit tobacco based on x-ray fluorescence and some advanced techniques. There was also a very important study done last year at NIST by Yi He and her colleagues. It was published in 2015 and some in NIST website, looking at lead and cadmium exposure in counterfeit cigarettes from China, and those are just soaked in cadmium. And it concluded, the study concluded that they have a higher risk profile. They present a greater public risk than legal product.

We have over two billion cigarettes coming into this country and also some went to Central American narcotics trafficking. Illegally made cigarettes are a unique product category. They have a different risk profile. We don't know what that is, especially when it's being done now in large quantity made with North American-grown tobacco.

There was a study in Australia looking at illicit-made cigarettes with Australian grown tobacco, and although chemical composition wasn't done, they did look at social function and a variety of factors and found that they had even worse health effects than regular cigarettes.

So what I would like to do is leave with two messages. One is that the FDA should look at illegally made cigarettes as a unique product category and determine the risk profile and the risks associated with these products, and also, of course,

is that the Agency will conduct its risk assessment proceedings in accordance with the White House bulletin on risk assessment and peer review. We've heard about the National Research Council and some other risk assessment frameworks, but it is the OMB one which is the governing document.

Thank you.

(Applause.)

DR. DRESLER: Thank you.

Now, I believe there's going to be another public speaker, but I'm not seeing the person here, so that's it. Thank you very much for the public comment session.

And so we'll move on to the first session, which is entitled Risk Assessment, but the first two speakers aren't here either.

(Laughter.)

DR. DRESLER: So we're going to start -- Robinan, are you -- is it possible that we could go ahead and start with you? Okay, thank you so much for being here, okay. So then the person who will lead us off will be Robinan Gentry from Ramboll Environ, and she will be speaking on State of the Science of Quantitative Risk Assessment for Application in Tobacco Submissions.

DR. GENTRY: Thank you.

Good morning. I'd like to thank the organizers for giving me the opportunity to speak this morning. For those of you who don't know me, I am a Ph.D. toxicologist and principal with Ramboll Environ. I'm approaching the 30-year mark working in the area of quantitative risk assessment. I spent the first half of my career working as a government subcontractor writing IRIS documents and ATSDR profiles and developing OSHA PELs, and our little group in Louisiana was obtained by Environ in 2001.

So today I'm going to talk about -- we've talked about a lot of the pieces, parts of quantitative risk assessment. I'm going to talk about the state of the science and how this is being used in the environmental situation and possibly how this same framework could be applied in tobacco submissions.

As a disclosure, the opinions in this presentation are mine. I'm very passionate about QRA, but funding for work related to these evaluations was provided by RAI Services Company.

So just to go over what I'll talk about today, first I'll talk about the state of the science for each of the steps for QRA, specifically as it's used to assess public health impact from constituents present in the environment and consumer products. And what I'll show is the similarity and applicability of these approaches for the environmental

situation to tobacco products via comparison, and also demonstrate the acceptance of these approaches across regulatory agencies and authoritative bodies for use in decision making.

So we've talked a lot about QRA. Some of these first slides are on the process in general. It's a scientific evidence-based analytical process where we can combine chemical and biological data to try and quantify the probability and potential impact of some defined risk. QRA has been noted by the National Research Council, the NAS, as a component for regulatory decision making, and it's used by multiple government and regulatory bodies to inform decisions about environmental, occupational, biological, and consumer product risk to human health.

So quickly, this is a four-step process; we've talked about some of the steps. I'll go through this quickly, but we go through these steps and combine them together to draw conclusions. The first step is the hazard identification step where we try to understand what adverse effects are associated with the chemical of concern; the exposure assessment where we understand how much of the chemical and by what route individuals are exposed; the toxicity assessment, which we understand how much of the chemical does it take to cause the

adverse biological effect; and finally, the risk characterization step when we put all these pieces together, what is the risk or probability of toxicity occurring in the exposed population.

So for the hazard identification step, what I have is a series of slides where we show the steps that could be used and what's currently used in the environmental situation and how that can be extended to tobacco product submissions.

For the environmental situation, during the hazard identification step, we look at the potential chemicals of concern as defined by EPA. The list of chemicals of concern are typically those representative of the classes of compounds expected to be present and those that have the appropriate toxicity data. With tobacco products, in turn, we have a list of harmful and potentially harmful constituents as defined by FDA, and this list of HPHCs, they're representative of the classes of compounds present and have the appropriate toxicity data as well.

Specifically, when we look at site-specific QRA, the list of chemicals of concern is typically driven by the chemicals expected to be present based on the activities and our operations at the site of interest. We don't test for every chemical that could possibly be there. There is a defined list

of sampling that is conducted for the chemicals expected to be present.

For tobacco products, we've talked about it being a large mixture of compounds, but we do have a list of 93 HPHCs that have been established by the FDA in 2012. Each HPHC is characterized as a carcinogen, a cardiovascular toxicant, a respiratory toxicant, or reproductive/developmental toxicant. There's also an abbreviated list for each type of product focused on those HPHCs for which testing and analytic methods are well established and widely available. They represent several different chemical classes and constitute a representative sample of the established list of 93.

For the second step of the process, which is the toxicity assessment, in the environmental situation we want to identify the exposure response information for the chemicals concerned. This process and the identification of these toxicity factors, which George talked about yesterday, it's based on a hierarchy of well-documented toxicity information, and in the environmental situation, the most sensitive endpoint is relied upon, assuming that if we use that as a toxicity factor, it would be protective for other endpoints.

In the situation of tobacco products, the same would be true for each HPHC under consideration. We would need to

understand the exposure response information. It could be based on a hierarchy of well-documented toxicity information, and again, the most sensitive endpoint could be relied upon in conducting the QRA.

George talked about this some. We have a database from the EPA that provides toxicity factors within the IRIS program. These factors are developed based on the most sensitive endpoint. For cancer, EPA provides estimates of extra lifetime cancer risk defined as the probability of developing cancer after a lifetime of continued exposure at a specified intake. We have those factors for both inhalation and oral exposure. And we have comparable values for noncancer, reference concentrations or reference doses or estimates typically based on animal data with an uncertainty spanning an order of magnitude of a daily intake for human populations, including sensitive subpopulations that's likely -- unlikely to result in adverse noncancer effects.

We have sources for these values, so they don't necessarily have to be developed. We have the IRIS program that provides many of these values; we have other regulatory agencies in the U.S. at the state level, such as CalEPA and TCEQ that provide these values, as well as ATSDR. In the absence of toxicity values provided by an authoritative body,

there are also studies that can be identified in the peer review literature for some of these compounds that are of the quality that they could be relied on for the development of a toxicity factor.

We noted that for a lot of these HPHCs, they're data poor. We do have approaches within QRA to try and understand the potential for toxicity for some of these compounds, specifically, route-to-route extrapolation. In the case of some of the HPHCs, we only have data available for one route of exposure. For example, we may only have oral data when the constituent is a concern for the inhalation route of exposure. We also have approaches where we rely on surrogate data for structurally similar compounds, or there's also the thresholds of toxicological concern approach, the TTC approach for substances with no toxicity data, but if we understand enough about the structure and the potential components of the structure that contribute to the toxicity, we could rely on these approaches to develop a toxicity value for these data-poor HPHCs.

So once we understand the potential hazard and develop a toxicity factor, of course, we have to understand the potential for exposure. In the environmental situation, we quantify the extent, frequency, and duration of exposure. It certainly

depends upon the media and the receptor when trying to understand the potential for exposure. And there are also steps where these estimates of exposure compare to background.

In the case of tobacco products, certainly we want to do the same thing: quantify the extent, frequency, and duration of exposure. This would be dependent upon the median receptor, product use, and the user, and there are some comparisons that could be done to exposure from other sources. For example, with smokeless products, looking at the constituents present, we may want to understand those concentrations as they relate to contributions from diet or drinking water.

So just to start with a basic equation in how we could possibly extend these environmental approaches to the tobacco situation, if you look at the EPA's RAGS Part F equation for exposure, we can start with understanding how we would estimate an exposure concentration. It certainly depends on the constituent concentration, the exposure frequency, the exposure duration, and averaging time.

A constituent concentration, of course, is a measured air concentration, or it can be estimated if we look at the oral situation from soil or ground water. So we have to estimate this for the evaluation of exposure from smoking. This could be substituted with a few additional parameters to transition

this equation, so we'd want to understand the HPHC yield for a particular constituent, understand the cigarettes smoked per day, and look at this as it relates to the inhalation rate.

So if we put these into the RAGS equation -- oops, sorry -- so we can see we have a basic exposure equation for combustible products that is consistent with the equations that we're currently using for inhalation at a Superfund site.

So for the final step of the process, risk characterization, we want to understand the adverse health impact due to the exposure that we've estimated combined with what we understand about toxicity. In the environmental situation, we typically look at the comparison of risks and hazard between baseline and a 5-year review incorporating the potential uncertainty and variability, and additivity is the current preferred approach for summation of risk or hazard when we are looking at mixtures of compounds in the environment.

Certainly, for tobacco products, we would want to do the same thing, understand the potential for adverse impact looking at the combination of toxicity and exposure. We want to, again, compare the risks and hazards in this situation between products for certain types of submissions, incorporating the potential variability/uncertainty, and additivity, again, is the current state of the science for looking at mixtures and

trying to understand the overall risk from that mixture.

I mentioned the additive model. It's documented in a lot of the environmental guidelines as a conservative approach, assuming independence of action by the constituents involved. There are a number of conservative assumptions that go into this process. For cancer, for carcinogens, the slope factor or the toxicity value for each carcinogenic constituent is based on the upper 95th percentile, so when we add things together, we are adding upper bounds as we understand the potential for health effects. The noncancer toxicity, it may be based on a NOAEL, but again, we're adding factors of uncertainty to bring those levels down to a concentration below where we expect any type of adverse effect.

Again, constituents, when we do this additive approach, constituents, a different weight of evidence regarding toxic potential are included in the sum. And certainly, factors for animal and human data are given equal weight.

So in the decision-making process, I have up here existing guidelines that are current on the use of additive approaches recommended by other regulatory agencies. Some of these are a little bit older, but there's been no guidance to replace those. These are still the state of the science that outline the simple addition of combining risk, which accounts for the

joint probability of the same individual developing cancer as a consequence of exposure to two or more carcinogens is considered appropriate. When evaluating predictive cancer risks from multiple contaminants, risk assessors should estimate the cancer risk for each subject, substance, and then sum these risks. Use of the dose additivity assumption is likely to produce estimates of health hazard that range from appropriate to somewhat conservative and which are therefore protective of public health. And the guidance recommends use of dose addition for determining the combined risk of the cumulative assessment group.

So here's just an example. This is not intended to be comprehensive or represent any constituent that you could be considering, but this is a standard table that's submitted for an assessment conducted at a Superfund site, so as you can see, and this is just focused on cancer, we would look at the baseline risk, which would be the results of looking at sampling at the initial investigation, and then look at a 5-year review, and we would look at the risk change between those different measurements and look at the total estimated risk to see what the differences are over that time period.

Similarly, this could be -- these approaches could be applied for a tobacco submission. Here, we're looking at

selective constituents. Certainly, again, this is not intended to be comprehensive. But looking at the estimated risk, we've got the mean and upper bound for Product 1 versus Product 2, and we can look at the total estimated risk and compare the products as a whole, but we can also look at the change within constituents. As you'll note here, some decreased, some increased, and so we look at -- some remain the same, so you can look at the overall difference between the products.

So overall, the QRA process, it gives us a series of steps to define the overall risk hazard; what I've shown here we can do it, we can use these approaches and these steps to look at differences in risk between products, as well as we have used them in the past to look at the differences between impacted areas versus regulatory measures.

In conclusion, the available science supports this approach. It allows for both the assessment of individual constituents as well as mixtures and understanding different questions of public health potentially associated with use of a new product. An approach could be developed consistent with approaches used by other regulatory agencies and authoritative bodies for decisions regarding public health effects. This approach provides a data-driven method; it's a way for us to use the most appropriate available science, which should

increase in confidence and decrease the uncertainty in the risk characterization for comparison across products.

And here's a few references I've relied upon, and I guess we'll wait for questions for the panel.

DR. DRESLER: Thank you.

(Applause.)

DR. DRESLER: Thank you. And look to me, I'll remind you of something I said yesterday. If you're interested in the slides, they can be obtained by e-mailing workshop.CTPOS@fda.hhs.gov. It's the same e-mail address that we used before to register. Give them a day or so because they're busy working on this, but you can get the slides, so if you want that reference list, that would be an easy way to get it.

Our next speaker is our first speaker, but is now our second speaker, so Vince Cogliano is from the U.S. EPA, and I believe 25 years -- how long at IARC prior? You were at EPA and then IARC and then --

DR. COGLIANO: Yeah.

DR. DRESLER: And then 20 years?

DR. COGLIANO: I'll figure that. I started at EPA in 1983.

DR. DRESLER: 1983, okay, for EPA.

DR. COGLIANO: And I had 8 wonderful years at IARC.

DR. DRESLER: So 8 years at IARC. Okay, so he will be presenting on Assessing Tobacco Products: Similarities to Assessing Chemicals in the Environment.

Dr. Cogliano.

DR. COGLIANO: Thank you very much, Carolyn. First, I'd like to apologize for not being here at the time you were expecting me to speak. I was thinking of 9:30, and I actually left the house in Virginia at 7 o'clock for what should have been 45 to 60 minutes, even on the -- at that time on the -- with the, you know, Google's directions. But there was an accident on the 14th Street Bridge, and that hour actually just took me up to the river, and then we got here pretty quickly. So that actually tells you something about risk and something about that it's reasonable to do upper bounds sometimes --

(Laughter.)

DR. COGLIANO: -- because I did allow myself a margin of safety, and I guess that got me here on time. So anyway, again, sorry for not being here earlier, but at least I'm here now.

So I'm going to talk about some guiding principles we use in assessing risks of chemicals in the environment at EPA, talk a little bit about mixtures of tobacco constituents, can

actually be compared, I think, to mixtures of chemicals in the environment. How we do make risk assessment for these chemicals and mixtures? I'll talk a little bit about systematic review, which is something that we're sweeping the field, I think, and we're certainly -- have really embroiled in this at EPA. And then a few other considerations.

So we all follow the risk assessment, risk management paradigm; you've probably seen this many times. It's been around since the Red Book in 1983 from the National Academy of Sciences. But we separate risk assessment into these four components that are in green: hazard identification, where we ask what are the health hazards that you might expect from some type of exposure to a particular chemical or mixture or other agent.

Once you know that there's a hazard, you want to do what's the dose-response assessment, how does -- what's the exposure-response relationship between exposure to that agent and the health risk that you just identified. And dose-response assessment has to account for extrapolation from high doses at which most chemicals are tested to lower doses that you see in the environment, animal to human extrapolation, high-dose, route-to-route, and any other kinds of differences.

Once you know that dose-response curve, then you want to

find out where you are in a dose-response curve, and so the exposure assessment tells you how people are coming in contact with the agent and how much of this agent they're exposed to so you can find out where you are in the dose-response assessment.

And the risk assessment, risk characterization is where you put together the hazard, the dose response, and the exposure to characterize the overall risk.

And at EPA, as probably at most other regulatory agencies, that's just one component of managing the risk, where you analyze and compare different options you might take, and you select the appropriate option, and you look at the risk as one of the inputs to that decision, but you also look at what your legal structure is, what are some of the political considerations, socioeconomic considerations, economics and technical consideration, is it actually even possible to reduce the risk to the extent you want.

So all of those things go into a decision; a lot of it is governed by the legal framework, and at EPA, we have several different legal statutes that govern how we regulate pesticides, other toxic substances, pollutants in the air, pollutants in the drinking water, and they do actually have different types of ways you balance all of these considerations.

Some guiding principles in how we assess risk. One statement I really like is from our cancer guidelines, so I think it applies to the way we look at all chemicals. "The primary goal is the protection of human health," so "[assessment procedures and default options] should be health protective." But at the same time, a critical analysis of the available information is the starting point for any evaluation we get. And we have to have assessments that are scientifically defensible and that is evaluated through independent expert peer review panels. I think that governs not just our cancer assessments but how we do any kinds of assessments at EPA. Now, that middle statement, critical analysis of the available information as a starting point, I'd say what kind of information is typically available for agents in the environment? So we're exposed to mixtures of chemicals in the environment; we're exposed to different mixtures throughout our lifetime.

The type of information you might actually want is toxicity data on mixtures in the environment, and sometimes these data are available; for example, for diesel exhaust, for PCBs, we do have studies that looked at mixtures in the environment and what the health outcomes are for people exposed in those ways.

But in many cases, the mixtures that are studied differ from the mixtures of interest; for example, diesel exhaust does -- is different if you're working in a garage where you're exposed to the diesel exhaust immediately afterwards, and it might be different if you're in a city where you're exposed to diesel exhaust that's weathered at some point and the particulates are gone one place, there's different types of half-life for the different types of compounds in the -- in what you're exposed to. Many times we want to look at toxicity data on combinations of constituents of these mixtures. We have that for a few combinations. Consider the millions of combinations people can be exposed to, usually binary combinations, and you think about if you just were exposed to 10, if there were only 10 chemicals in the world, you'd have 100 binary combinations, and then you got to think about am I being exposed to these at the same dose, or am I being exposed to one at 10 times higher than the other, and you think about the almost infinite number of possibilities of combinations that you can think about, and so it's a small fraction of what we're really interested in.

So most of what we do, when we're assessing mixtures in the environment, is to look at toxicity data on individual constituents. And we do have those toxicity data on hundreds

of common constituents, and we often add them. And you've probably heard already about dose addition and response addition, but they're added in some way, and one of the difficulties is this ignores interactions which you would get if you were looking at binary combinations or at whole mixtures like diesel or PCBs that are studied in the environment.

But this is what we often do, and this is also, I think, a byproduct of the regulatory structures that we've had at EPA since its inception. The Clean Air Act from 1970, the Safe Drinking Water Act says that the EPA should set standards that are safe for people to consume for individual chemicals, and so we have, you know, lists of dozens or hundreds of chemicals and what the safe level of exposure in the air and the drinking water or in soil or other media would be.

So our regulatory structures are dictated a lot by our laws, and our laws say to put standards on individual chemicals, so that's why we have those test data on individual chemicals, and that's why we have all these toxicity values on individual constituents.

Now, one of the messages is, you know, tobacco products involve mixtures of many constituents, and so do chemicals in the environment, and the Center for Tobacco Products has a list of harmful and potentially harmful constituents, and they take

them from various sources of agencies that evaluate individual chemicals, things that are known or possible or potential carcinogens by IARC, EPA, NTP; respiratory and cardiac defects identified by EPA, ATSDR; reproductive and developmental toxicants by CalEPA; and some other things, like measures of addiction or some things that are banned in food. But these are generally individual compounds that have generated these lists, and it's because they come from agencies that are, for many decades have evaluated individual compounds, and that's the way we regulate our environmental mixtures.

I'll say that at EPA in the IRIS program that I'm in, we're evaluating a lot of chemicals that are also tobacco constituents, so we have 20-some ongoing assessments that are listed in the column on the left, and the second column, if there's a green checkmark, these are also on the Center for Tobacco Products list, so about half of them.

And we have a list of 15 future assessments we'll be taking up in the future, and of those, about half of them, again, are also on the Center for Tobacco Products list. So you've got a lot of information from EPA on many of the chemicals that you need to look at and so -- and that's also the case for IARC or for some other agencies. So there's really a good commonality between the evidence base that we're

all looking at.

So go back to the risk assessment and risk management paradigm. I'm going to go through what we do at EPA; I think you've seen parts of it, I've seen parts in the previous talk. I'm just going to present, I think, some of the same information in a somewhat different format, but I think it's good to hear something from multiple perspectives because I think that there's a lot of commonalities between different -- the way different agencies look at these kinds of exposures and take actions on them.

So hazard identification, one of the first things we do for cancer is we consider the animal, the human, and the mechanistic data and classify agents into various categories as carcinogenic or likely to be carcinogenic, or we say we have suggestive evidence or inadequate information. And sometimes we'll say that we have substances not likely to be carcinogenic.

This is very similar to what the International Agency for Research on Cancer does; they have slightly different guidelines, but they overlap to a good extent. Other agencies also have different kinds of classification schemes. For noncancer hazards, we don't have a classification scheme at EPA; we generally have a narrative discussion about whether we

have enough information to lead to the derivation of some kind of a toxicity value that you want to, you know, keep exposures below.

Now, for dose-response assessment, we look at the best studies we have. Those are generally going to be studies that look at a wide range of doses, so you can see a wide range of responses and fit a dose-response curve, generally multiple exposure levels so that you can get some information about the shape of the dose-response curve. The more subjects, either animals or humans, the better you -- you know, the more precise you can make your estimates, the better you can understand that dose-response curve.

You look for the types of exposures that are closer to what, at EPA, you would get in the environment, so we would generally prefer a study that's at lower exposure levels than high-dose studies at the maximum tolerated dose, and you look at all the studies; you decide which ones are going to give you the best estimates, you fit a dose-response curve, and then because we take it generally down to some benchmark level, generally 10%, but it can be -- it could be something different depending on what levels we can estimate well; if we have a large human study, we often go down to 5% or 1%. But you go down to that level to find the dose associated with that level.

We take a lower bound on that dose because, again, it would be health protective to account for the experimental variation that if you ran that experiment again, those points would bounce up and down a little bit, and so that lower bound, 95% confidence bound, gives you some protection against experimental variation, and then we generally take a linear extrapolation down as far as we need to go to estimate the exposures in the environment that we're interested in.

Sometimes this is criticized as going all the way down to zero, and we're saying that one molecule can cause cancer, and that might be a nice theoretical argument to have, but in 25 years or so at EPA, I've never seen this, evaluate an environmental exposure where we're exposed to one molecule, even when we're taking a very potent chemical like dioxin and we're looking at picograms per kilogram per day as a dose; that still is billions and billions of molecules.

So we generally don't have to go all the way down to zero. We go down as far as we need to go basically to estimate environmental exposures. For noncancer hazards, we start with the same general approach. We look at the studies we have, we get the best studies based on there being multiple data points, lower doses, systems respond like humans and so on. We fit a dose-response curve, we take it down to some kind of a low

dose, a low response level like 10%, take a lower bound on that for experimental variability, and we drop that down to a dose that's called really the point of departure, but -- and then we apply uncertainty factors because the experimental situation is not necessarily the same as the human exposure situation we're looking at.

So we look at human variation, how sensitive might some people be compared to the average human. We extrapolate from animals to humans. In many cases now we're using physiologically based pharmacokinetic models to do that, to look at various differences in metabolism and exposure rates.

Sometimes we don't have a chronic study. At EPA, we're generally looking at longer-term chronic exposures, but sometimes, if we don't have a longer-term chronic animal study, we'll use a shorter-term study, maybe a 3-month study, a 6-month study, and we say -- we have another uncertainty factor because the animal exposed for chronic period might show the same response at a lower dose level compared with the 3- to 6-month exposure period that we might use. Sometimes we don't have something that we can model, and we just take the lowest exposed lowest level where you do see some kind of response, and we'll apply uncertainty factors to try to get down to a level that would have no observable risk or something lower

than that. And sometimes we don't have a very complete database, so we might apply an uncertainty factor for that.

So we apply these uncertainty factors so we get the reference dose for oral exposure or the reference concentration for human, for inhalation exposures, and that's the dose level that we would like to see environmental exposures be below.

So how would we characterize this? So you'd go out, estimate exposures in a particular environmental scenario; it might be a national regulation like a drinking water standard. We assume people are going to drink a certain number of liters of water every day for their whole lifetime. When you figure that out, it might be at a hazardous waste site, a Superfund site where you figure people might be exposed based on their distance, based on the prevailing geological and weather patterns, what they might be exposed to, the concentrations, how many hours a day, how many days a year, how many years, and you put it into the exposure equation that you just saw in Robinan's talk, and you estimate some kind of a concentration; for example, I did inhalation for this. You take the unit risk estimates; that's that low dose slope for cancer. So for multiple chemicals you would have at a site, you would take the unit risks that you calculate, and I'm just making these up with simple numbers so it's easy to see how they're multiplied.

You multiply them by the concentration, and you get a risk from each individual chemical, and then you add them up, and that's your total risk.

And the risk manager or the person setting a national drinking water standard, the person evaluating the level of cleanup at a waste site, the person you're evaluating, the number of fish meals you might be able to consume with a hazardous fish contaminant, you decide is this an acceptable risk or not, and if it's not, you figure out how to reduce the exposure to get the risk to some acceptable level.

For noncancer hazards, it's something fairly similar. You would go sample the concentrations and estimate them with the exposure equation. You would look at the referenced concentrations that were calculated onto that dose-response extrapolation, and you would take the concentration that you see over the referenced concentration which you might think of as an allowable concentration, and you see what fraction of that allowable concentration are people actually being exposed to. And you just do the division, and for multiple chemicals at a site or at some kind of exposure, you add up these quotients, and you get what's called a hazard index, and once again, you have to decide whether that's an acceptable hazard index.

Typically, we want to see the hazard index less than 1, but again, these are decisions that take into account not just risk but technological feasibility, legal considerations, economic considerations, and these are decisions made by risk managers in the various program offices at EPA.

By the way, our organization is located in the Office of Research and Development so that we can support people in the office that does air regulations or drinking water standards or hazardous waste cleanups or other toxic substance regulation under TSCA so that we really support all of these equally and so we're not really associated with any one program, we're centrally located in EPA.

Okay, we have this resource of hundreds of chemicals in the environment that we have evaluated, it's called the IRIS database, and in that you'll find classification, an oral slope factor and an inhalation unit risk if we determine it's a known or likely carcinogen. And for noncancer effects, an oral reference dose or an inhalation reference concentration. For those effects, those are the levels we feel that it's acceptable to be exposed if we're below those levels.

For every chemical you want to have all of those. In past years, we had separate groups doing the cancer evaluations and the noncancer evaluations, they got somewhat out of sync. Some

chemicals, we have cancer but not the noncancer evaluations; in some cases we have oral but not inhalation.

One of the things we're trying to do in the Integrated Risk Information System is to do a comprehensive evaluation of cancer/noncancer effects and all relevant routes of environmental exposure. So, many cases you'll find all of these, but many cases you won't. But we do have hundreds of chemicals in the environment; it's the first source that EPA risk assessors consult.

Okay, now we're doing our evaluations by systematic review. It's basically an approach that promotes objectivity and transparency in syntheses of published research. So what we want to do is identify pertinent studies, and what we do there is we really document the way we do literature searches, the way we decide that a reference is pertinent to the evaluation or not pertinent to the evaluation. We put those criteria out before we do the literature search, and we do the literature search, and we make sure that we are identifying all the pertinent studies.

We also do a lot of public engagement in our evaluation, so our assessments are peer reviewed. We also get public comment, and we can ask did we miss any studies, and if we did, there's ways to put them into our assessment before it goes to

peer review. Then we want to focus not on every study but the higher quality studies, so we develop considerations for what makes a study higher quality or not.

It's something that actually is chemical specific and disease specific. As we've been doing many of our evaluations, for example, our evaluations of phthalates, something that surprised me is you really want urine measurements of phthalates, not blood measurements of phthalates. I would've thought any biological measurement would be good, but just the way it -- it reacts with other constituents in the blood, you really want urine measurement. So for each different agent you're looking at, you have to decide what makes -- what's the best measures of exposure, what are the best measures of outcome for each disease, what makes the particular studies higher or lower quality. So we evaluate them and try to focus on the high and the medium quality studies and to eliminate the studies that are critically deficient in some respect.

Then we integrate the evidence; that's where we come up with the cancer weight of evidence. Or for noncancer effects, we decide whether there is enough data -- there are enough data to develop reference doses.

Then we select studies for deriving toxicity values for every health effect that we determine is associated with the

chemical, and that's because the studies that might be good for evaluating dose response might not be -- the studies that were good for evaluating hazard might not always be good for evaluating dose response. You might, for example, have an occupational study where you know people are exposed or not exposed, and you can get an idea that there is a hazard among people who are exposed, but if you don't have a good quantitative measure of exposure, you can't do a good dose response for that. So we go through the studies again to try to get the best ones for toxicity values, and then we derive them.

We also do a lot of planning. We found that several assessments that had been ongoing, as we get to the peer review stage or the public comment stage, we find things that we missed. So now we're putting a big emphasis on scoping and problem formulation. We have a public meeting with our problem formulation to make sure we've gotten all of the right issues. We write the protocols for how we're going to do the literature searches, how we're going to do the study quality evaluations, and then we go do that.

Systematic review is great, I think, particularly for our hazard identification, but there's a lot of questions that we deal with at EPA that lie outside the realm of systematic

review, and that's key science issues, our particular health effects in animals relevant to humans, how do we analyze mechanistic information and what it might say about who is more susceptible or less susceptible or what the shape of the dose-response curve might be, how we model these exposure-response relationships and how we would select toxicity values when we have multiple type -- multiple options for that.

Just a few other considerations about what makes environmental risk assessment so complex. Toxicity data are really complex, and sometimes only there's a very small number of specialists who can participate effectively in these discussions. Sometimes you really need to be an expert in how a particular mechanism of carcinogenesis works to really be able to fully debate those data, and so we often have a small pool of people who are looking at these things.

And this is very different from 20 years ago when you would look at epidemiology studies or animal studies and say, yeah, look at this, you do have a dose response. Now we're trying to figure out why do we have a dose response, and what does that mean, and you have a smaller number of people who can really effectively debate those.

Actions are often debated based on reasonable inferences from data rather than on direct observation of adverse health

effects in a population. We don't want to be relying always on epidemiology studies even though these are, in some ways, the best information to tell us whether we have a risk. We would like to be able to test these in animals or in in vitro systems ahead of time, and in that case, we're really making inferences from indirect data to be protective of human health. Human populations vary in their level of exposure and biological susceptibility and certainly, also, in just a variation, a variability of different types -- I'm too short to see that, sorry.

(Laughter.)

DR. COGLIANO: I'm close to the end anyway. And we're making risk levels that might not be observable. At EPA, we generally look at risks of 1 in 10,000 to 1 in 1,000,000, and you have to have a very, very large study to be able to identify a risk like that with confidence.

Sometimes it's difficult to observe. This is like a Superfund site; I think it's kind of interesting. Suppose you have, right next to an industrial plant or a waste site, a small population with a high individual risk, and then you have an intermediate, a larger area, and then a much larger area with a larger population, but a smaller risk level. You can multiply these things out if they're linear, and you can find

out that the population that has the highest risk has really the smallest numbers of expected cases, and most of your cases are actually in people who are exposed at quite a low level. And this is actually a paradox that when you're trying to take remedial actions, the populations at highest risks are not really where you're going to get the largest risk reduction, and so this is one of the things that makes our environmental decisions so complex.

So, in summary, I think we all deal with mixtures, but we have data on individual constituents, and that's something that's going to be with us for some time, I think, until we learn how to use a lot of high though-put data; then maybe we can actually start testing the mixtures and get a lot of information on that, but right now we're dealing with individual constituents.

Pragmatic approaches have proven useful. As we understand more and more about mechanisms, our assessments become more and more complex, but that's not to say that the basic assessments that we do with these basic methods haven't been very useful in protecting public health.

And systematic review is becoming widely adopted, but it also doesn't address some of the pressing questions we have about mechanisms, about susceptibility, and so we have to

continue with a lot scientific judgment.

So I want to thank you for your attention.

(Applause.)

DR. DRESLER: Thank you, Vince.

Okay, our next speaker is Dr. Peter Shields from Ohio State University, and he will be speaking on Risk Assessment for Tobacco Products: Examples and Challenges. Peter.

DR. SHIELDS: Good morning. I think I'm going to get the slides, or do I need to do that?

DR. DRESLER: No, they'll --

DR. SHIELDS: So I think this is going to be a very different type of talk than you've been hearing about. It's going to be very tobacco focused. I and a group of collaborators have been thinking about risk assessment for tobacco for 15, 20 years, and so I'm going to be taking a very academic approach in making this process extremely complicated for a process that is already very complicated. I'm going to talk about a framework for tobacco product evaluation that we developed. Of course, I'm going to talk about weight of evidence and causation, but contrast critical issues that I'm not sure I heard many or anyone talk about in the last day or so about individual versus population risk assessment. And I'm going to give you some examples of what's out there now,

although I'm not going to talk about one of the coolest things that are being studied now about low nicotine yield cigarettes.

So, first of all, following a process in the mid-2000s under an NCI contract, a group of investigators got together to think about how do we evaluate a modified risk tobacco product. So there were several on the market, emerging ones on the market; this was before the FDA had the congressional authority. And a group of experts in numerous workshops came up with this framework, and we issued it as a report. Then FDA gets the authority to regulate products, and so we kind of modernized this and published it last year. Micah Berman, who is in the audience, is the first author of this paper.

And so we basically wanted to provide a framework for evaluating all tobacco products, not just modified risk tobacco products and product standards. We were thinking back then about how do you improve the public health related to tobacco products and prevent a worsening of it. We wanted to prevent unwarranted health claims, minimize consumer misperception, provide an early warning for unintended consequences, and all of these things -- and identify research gaps -- all of these things don't, aren't taking into account for your typical risk assessment. You don't think about perceptions or behavior, or how drinking water interacts with the person, and how if you

might change benzene in drinking water, whether or not someone's going to drink more water or shower more often. I mean, that's not part of the process, but it is in tobacco regulation.

So we have multiple sort of tables and figures that outline the types of testing that we'd want to see before modified risk tobacco product hits the market, or is not substantially equivalent, or product design, and that includes both your laboratory testing, but initially even some human testing, and not just toxicology but the perception, biomarkers, what's happening with how the person actually uses the product.

And then if you're actually thinking about claims and product standards, then there's additional analysis and postmarket activities, and then we recognize that the marketplace changes, so new products come on the market, and what might be modified risk today is no longer an acceptable risk because there's a new product that might have less risk. So you have to go back and start the whole thing all over again.

So I'm not a weight of evidence expert, and so we've heard a lot of good stuff about these sort of things already about quantitative weighting, systematic reviews, but remember that

in the tobacco world, it's very disparate data, so it's not just -- and I don't mean this just to minimize, but cell culture, animal, and people, but it's behavior and how they're going to do it and what they perceive and how that might affect their addictiveness of a product and that sort of thing.

And we have to think about the unintended consequences, so the analogy will be, yes, you want to regulate a certain metal in a fish, and you tell people that there's metal in a fish, and if people eat less fish, that may not be good for the public health, okay. So we have to think about that in our weight of evidence for tobacco. And you have to distinguish between effects that you can measure in clinically meaningful effects, so that's also very important. And then, of course, you come up with a qualitative final assessment.

So just for historical purposes, you know, the weight of evidence type reviews for tobacco, 1964 and the first Surgeon General's report linking smoking to health. We talk about the Bradford-Hill criteria. Actually, this report came out before Bradford-Hill, you know, published his paper, and so people in the tobacco control community get all angry when they talk about Bradford-Hill and not the Surgeon General's report.

And then in 2004 the Surgeon General's report updated the thinking that was done back in 1964, so there's actually a

chapter there on introductions and causal approach, and it approaches the causal inference. And the reason why it's important is because this is what tobacco control people think about and their priorities. I just went backwards for some reason. Okay.

So here's a fundamental problem we have. You could have a product that reduces or a product design change that reduces an individual's risk, okay, which is defined by intensity and duration of use, but you may mess up the population. And the FDA's mandate by the congressional law is decisions that are appropriate for the protection of public health, so if you have a product that might reduce an individual's exposure but then there's more people using this dangerous product even though it might be less dangerous than other things, you may have more cancer, you may have more heart disease. So this has to be thought about in a risk assessment, so not just less exposure good. There's less exposure, as we know, as I'll come back to with some of the low tar cigarettes ending up making a more dangerous product.

And then you have to think about what's the comparator. So are you talking about risk compared to smokers, and what's the impact on maybe delayed or decreased quitting, or are you talking about former smokers, because even former smokers get,

you know, heart disease and lung cancer even though they've successfully quit. Are we talking about the never smokers, the kids, okay? And then there's also the important group that's not looked at is dual and poly-users, so a lot of tobacco users out there are actually smoking using smokeless tobacco. The number of smokers now using electronic cigarettes are huge. And so that all gets factored in somehow.

So individual disease risk is toxin exposure, so the product, constituents, formulation, design, and delivery, but the use is not so simple. We can think about what an average person showers or drinks water, but this depends on the individual. There is susceptibility to becoming addicted, there's susceptibility to marketing and product appeal, consumer perception, what their health risk knowledge is, and maybe even some genetics.

And that's, again, interactive with what's actually out there in the marketplace. The price and accessibility, availability of alternative tobacco products, marketing, and the environment. Do you live in a household of smokers, of smokeless tobacco users, or your friends are all using e-cigs and that sort of thing? So again, from an academic perspective, this gets complicated really fast.

And then for population risk and population harm, you have

all those individual stuff, but then you bring in the prevalence of use. So how abuse liability, how addictive is the product going to be? What is the typical consumer perceiving about this? And then the price and accessibility, because they all impact on uptake, continued use, and delayed quitting. So the appropriate protection of public harm or population harm really is a lot in this box down here. So when you think about this, okay, if you have, for the individual cancer risk in a user, duration of use, the longer you smoke, for example, the higher your risk. If you quit smoking, you kind of freeze that risk, okay? But let's say you have a product standard or modified risk tobacco product that still -- it's a combustible product, so there's still more risk than none.

You may have some improvement, but what if you have some people who are dual users, but they reduce their smoking enough for this other product and they have a reduced risk, but you can also have an unintended consequence where you may even have higher levels of exposure. This is all hypothetical, okay, but it makes it complicated; it's not just one dose-response curve that people have to work on.

And then all of these things, again, get affected by not just the product, but the age when they make a change, so the

product may -- or the product design may impact someone more if they change over the age of 20 versus the age of 60 when they've had 40 years of smoking and their perception and their nicotine dependence and who they are. And then again, as I said, very important to think about what also is in the marketplace. So the population, okay, so the risk assessment is different in tobacco because you have to account for the delayed quitting, new users, relapse. We have a very heterogeneous population, so by race, you know, you'll see more menthol smokers; different regions of the country will be prone to use different products. You have to think about different disease outcome, so typical risk assessment, if it causes any type of cancer, it's a cancer-causing agent, and it gets regulated, okay?

But here we're talking about potentially trading, okay, lung cancer, for example, with a product design change that might reduce lung cancer but may increase heart disease risk, okay, and depending on how much you do, that may be progress. That may be an acceptable risk; it may not. So you have to think about also overall morbidity and mortality. And then I'll come back to this reverse dose-response relationship, but we make the assumption that if something more causes more risk, then less will cause less risk, but maybe not; maybe you can't

go backwards that easily.

So here, on a population level, cancer incidence in a population and the change in cancer incidence over time, if you just don't have the product in the marketplace that can potentially reduce risk, that's your level. But if you can get more people to quit, okay, then you'll have complete cessation, and you'll have cancer incidences go down over time. And if you have that product on the market that actually is a reduced risk product, then depending on how much lesser smoking you have, you may reduce risk, but maybe we can get people to just smoke less cigarettes and reduce their risk.

And then, finally, there's the unintended consequences that we have to worry about, okay. And all of this gets affected by, again, the late cessation, increased initiation, reuptake in former smokers, as well as the product and who the person is. And again, there's the complex marketplace.

So this is the reverse dose-response relationship that I talk about. So disease risk, okay, increases with more lifetime exposure. So a never smoker has a very low, for example, lung cancer disease risk, and a smoker has a very high lung cancer disease risk. Sorry, I'm just -- so the question is, is if you take a 30-year smoker and you've now reduced their exposure, are they going to drop down at all, and maybe

they won't. So we know that if they completely quit depending on when they completely quit, quitting at any time in your life is a good thing, but how much will actually be reduced and what's clinically meaningful if we reduce -- you know, we surgically reset the tobacco-specific nitrosamine NNK in cigarette smoke, is that going to be a clinically meaningful reduction? And from a policy perspective, that's going to have to be decided, is how much risk is measurable and acceptable, and if it's not measurable, then it may be still acceptable from a policy perspective.

So this complicated slide is here to say that there are studies that will give us a sense of how much exposure reduction you need for smokers to reduce all-cause mortality or lung cancer risk, and it's actually pretty complicated. We don't have a good sense of what the dose-response relationship is. So there are studies that look at smokers in a cohort study where they've reduced their smoking from beginning to end by about 75%, which is a lot of reduction in smoking. And the data is really mixed about whether or not mortality goes down or lung cancer goes down. And so is that the benchmark, that you want a reduction in exposure at least that much or something less? That's really a policy decision.

Then we've got biomarkers, so we talked a bit yesterday

about the biomarkers, and as you get closer to predicting disease, our biomarkers become less precise, and in fact, we don't have good biomarkers in this area of biologically affected dose of biomarkers of harm; we're focused on mostly the external exposure assessment and internal doses and even less so on target tissue. So we have biomarkers that are critical to understand exposure, but they're not that good right now, but it is the best that we have.

So this complicated slide is to remind me to mention that it's not just one critical biomarker; it's not like just cotinine as a nicotine metabolite, but as you look at a panel of biomarkers, you get different results, and so some things are statistically significantly predictive of lung cancer risk and some aren't, and so you need the battery to sort of screen across different risks.

So I put in two slides yesterday because there was a lot of discussion on the omics, okay. So I want to try to give you my perspective of why we're doing omics studies in smokers. So this is a paper that we just published the beginning of this year for metabolomics in smokers, and here is -- so we get a number, it's untargeted metabolomics, so you get a mass spec, and you get a bunch of peaks, and you look at what peaks are different between, in this case, heavy and light smokers, and

you see that you could separate them out somewhat. So the green is the light smokers, the reds are the heavy smokers, and you could imagine that if the heavy smokers are the greater lung cancer risk, if you have a modified product and you take these red boxes and you move the red -- they're not boxes, triangles -- and move them towards the green, you can infer that maybe you're reducing the risk because you're making a heavy smoker more like a light smoker. And you can model these things, and that's the theory, okay, but that's why we're doing that, to show you an example that's -- oh.

And then in this paper, we actually were able to do this based on smoking level, but also race and gender, so who the person is, whether you're doing it -- so we had a blood test before a cigarette, and then a few minutes after the cigarette, and that's different, of course, it's by their genetics and nicotine metabolic ratio. And then we published another paper just recently on metabolomics for the menthol smokers versus the non-menthol smokers.

So to give you another example, and this is extremely preliminary data, don't even ask me about the statistics, basically it's a pilot study where we're doing bronchoscopies on never smokers, e-cig users, and smokers, and all we want to do is just to make sure that we're handling the samples right

and getting quality data. But we did take a quick look, and I wanted to show you, this is very preliminary, but you really have, in an unsupervised analysis, very clear separation for gene expression of the smokers, e-cig users, and never smokers.

Now, you have three different groups, but we have no idea whether or not this group is more like a smoker or a never smoker. Does this give us any information about toxicity? You know, I'm going to come back to that. But you can imagine that if you wanted to regulate an e-cig and use this type of omics data, if certain voltages of the e-cig moves these green folks closer to the blue folks, that might, you know, play into a policy decision on the regulation of e-cigs.

So I want to give you some, a few real-world examples. So just recently FDA submitted to OMB a proposal to regulate N-nitroso or nicotine NNN in smokeless tobacco products. No one other than the FDA and OMB knows what's in it, but it's certainly obviously something that's in play, but I will tell you about some things that have been written. So a couple of years ago Irina Stepanov, Dorothy Hatsukami, and others worked for the Campaign for Tobacco-Free Kids and submitted to the FDA a request to consider regulating NNN and NNK in smokeless tobacco. They published a summary of that just this year, and basically they said we know NNN and NNK is bad; we know that

you can effect NNAL levels in people based on the levels of NNK in smokeless tobacco, so less in their mind was better.

This was not a weight of evidence review; it's just an evidence summary, and they basically said that this is a good idea. So it was basically a summary without a recommendation, but more of a request. But a lot of work has been done on smokeless tobacco, and I want to give you a couple of very different concepts. So this was published by Yusuf and Connolly several years ago, and they followed, which I'm going to show you in a minute, for smoking, what they did is they said let's look at the chemicals in smokeless tobacco; let's look at what EPA says about those chemicals and which ones are more potent and less potent and use it as a cancer risk end disease and the slope factors and that sort of thing, and can we classify different products based on at least what the EPA thinks is bad in the environment? And the paper then went as far as to say if we use an EPA benchmark, all of these products are unacceptably high risk, but we know that smokeless tobacco causes oral cavity cancer, at least the ones that we're selling here in the United States primarily, so I thought that was kind of silly, but their analysis points out that, you know, we tend to focus on the tobacco-specific nitrosamines in smokeless tobacco, but their point is all this other stuff should also be

included for a total assessment of comparing one product to another.

You know, there's a lot of reasons why this is imperfect, but at least, you know, there are folks that are thinking about this complex mixture in tobacco itself and following the EPA, rightly or wrongly.

We extended that sort of analysis based on data that we had on 18 smokeless tobacco products, or 19 products using 18 chemical constituents. We found statistically significant differences in these products, so the conventional products had higher tobacco-specific nitrosamines, ammonia, benzopyrene, cadmium, nickel, nicotine, and nitrate, whereas the products, the lower snus-like products had higher arsenic, lead, and chromium. Statistical significance, so what does that mean?

Okay. So we also, then -- we did this principal component analysis, and we said, okay, if one assumes that the snus-like products are lower risk, we could clearly separate these out. What if you started mandating different constituent levels in smokeless tobacco products to make this group approach this group more? Okay, so that's the thinking. And you could incorporate something like that into a policy decision.

But we also went and looked at those cancer risk indices and smoke and slope factors, and said, well, why don't we add

up these things and see what their risk is based on, sort of, an EPA concept or a California concept, and you can see that these products actually get closer together, and that might be a good thing or a bad thing, but in terms of meaningful, you get a threefold higher risk for the conventional products over the snus-like products, but that's threefold using the one-in-a-million standard, okay? So is that meaningful at all? And these are challenges that we have in risk assessment.

Moving to cigarette smoke, this was published 2003 by Fowles and Dybing where they did the same thing; they said let's look at all the constituents in cigarette smoke, let's see what the cancer risk indices for each of these -- and the analysis was very interesting because what they said is that we tend to focus, in tobacco control, on the TSNAs in benzopyrene, but if you look at the EPA-predicted risks, most of the risks from smoking comes from the aldehydes and small organics, the metals and benzene, okay?

And my point is, you know, are we really looking at the right things? I mean, we need to be thinking that through, and you're looking at individual constituents, how are you impacting the other constituents? Lots of limitations to this type of approach, including use of the smoking machine, which I'm going to mention in a second.

Steve Hecht published a paper in 2014 calling for reduction of NNK in tobacco, and basically, his standard was go to the lowest that's on the market right now, and I'm not sure that's a good policy decision, but anyway that's what he was calling for. Interestingly, one of the quotes that I wanted to take from here was that he said, "All data summarized here strongly indicate that a decrease in lung, oral cavity, and esophageal cancer incidence among long-term smokers of these modified cigarettes" -- the low NNK cigarettes -- "will occur, but predicting the extent or the timing of that decrease would be speculative because there are too many other variables." But, in fact, that's what the FDA's required to do because they have to think about what's appropriate for the protection of public health.

So the last thing I'm going to mention to you about is an effort that a group of us are working on right now to look at filter ventilation with a weight of evidence review, that we are using the words, and I'm not going to go through the method itself, highly suggestive evidence for causation, filter ventilation increasing lung adenocarcinoma.

And just so that we're all on the same page, what is filter ventilation? So cigarettes have filters, and there are these holes in the filter that allow for smoke to be diluted

with air, and it lowers machine measured tar yields, which at one point in time was thought to be a good thing.

Lots of different ways of lowering machine tar yields, but filter dilution is a critical way, and on the market today, although this date is about 10 years ago, the filter ventilation levels in cigarettes range from 0% up to 81%, so this is something that has to be -- so why are -- these cigarettes have such a wide range, and is there a problem with the filter ventilation? Clearly, at this point, there's no benefit to having filter ventilation from a health risk.

So we decided to follow the work issued in 2014 from the Surgeon General's report that concluded that lower tar yield cigarettes are more dangerous. They said "the evidence is sufficient to conclude that the risk of developing" -- oops, sorry -- "adenocarcinoma of the lung" -- one of the lung cancer subtypes -- "has increased since the 1960s." At that same time, they said that "the evidence shows that the decline of squamous cell carcinoma follows the trend of declining smoking prevalence."

They said that "the evidence is sufficient to conclude that the increased risk of adenocarcinoma of the lung in smokers results from changes in the design and composition of cigarettes since the 1950s," and "the evidence is not

sufficient to specify which design changes are responsible for the increased risk of adenocarcinoma, but there is suggestive evidence that ventilated filters and increased levels of tobacco-specific nitrosamines have played a role," so we decided we're going to think a little bit more about this. And it's not one or the other; the FDA can regulate both of them and many more things.

So this is the way the epidemiology's been going. Smoking has been going down, you would expect less lung cancer incidence, so that's what you have seen over time for squamous cell cancers, but not for adenocarcinomas. And since people's genetics don't change or how cancer causes in people don't change, it was reasonable to assume, which is what the Surgeon General's report did, is that it's the change in cigarettes over time.

And so if you look at a variety of different types of data, more modern cigarettes tend to be more risky for adenocarcinoma, so 19-fold for more recent studies compared to 4.6, studies from older cohorts.

So our weight of evidence puts this mechanism together. We know that filter ventilation reduces the time -- it increases the burn time but reduces the temperature burn, so you get less complete combustion, which ends up with an

increase on a per milligram of nicotine or tar basis, increased mutagens, changes in proportion of chemicals and increased tobacco-specific nitrosamines, and at the same time -- so that's in the laboratory -- what's happening in people, okay, their behavior as you have more filter ventilation, they take bigger puffs, they take more smoke into their lung, they're taking more smoke of this chemical, of the chemicals that are there in higher proportion and with increased mutagens.

They may also be inhaling larger particles that can sediment in deeper areas. They may be doing deeper inhalation, the data is not as clear as it should be, and people also smoke more cigarettes per day. The smoke has a better chance of getting to parts of the lung that are more sensitive, the cells that are more sensitive to the tobacco-specific nitrosamines, and hence, there's more lung adenocarcinoma. So that's the way we put the weight of evidence together.

Data, of course, is not perfect. You know, we don't live in a perfect world, but we think that that's sufficient for the FDA to include or to make some decisions about regulating filter ventilation.

I'm going to actually skip over that in the interest of time.

Very important about the smoking machine: We all know, we

heard this yesterday several times, that it's not predictive of actual exposure in people, but the reason why it's not predictive for exposure in people is because of filter ventilation. And I would submit that any change that one has to do, the laboratory stuff is important for mechanism, but we have to remember that with filter ventilation in the cigarette, the cigarette remains elastic enough that we really don't know what exposure is going on linking the laboratory work to the humans, and if you want to think about regulating any of the harmful chemicals, you know, I wonder whether you could do that without regulating filter ventilation at the same time because you're not going to know what the actual exposure is to individuals.

Other expected outcomes of such regulation is we don't think that there would be an adverse effect on smoking behavior or increase in toxic exposures. Essentially, the smoke becomes harsher; people would probably not smoke more. We don't think they'll be more addictive even though you may have a higher nicotine delivery. These are our expected outcomes, our hypotheses. We think it will reduce the cigarette appeal because of the increased harshness of the smoke, and it will probably facilitate transitions to other things like nicotine replacement therapy, electronic cigarettes, or quitting. But,

in fact, there's a lot of data on ventilation causing issues, causing changes, changes in behavior. There's actually no studies on switching people to unventilated cigarettes.

So that's a very rich research agenda right now is if you actually switch people to -- you take out the filter ventilation, what's the impact on toxicant exposure, what's the impact on the addictiveness, and that could be tested both in the lab as well as in people. How does the variation in cigarette ventilation along with changes in packaging -- so you don't imply that these are less harmful cigarettes/more harmful cigarettes based on packaging -- how does that affect their use? And then how does the regulation get moderated by the availability of other alternative nicotine delivery systems.

So, just to summarize, a weight of evidence integration requires consideration of lots of different types of scientific data. How to weight the studies are a big open question in terms of human studies, which types of human studies, clinical trials versus epidemiology, the strengths and limitations of both and how you weigh those more than your lab data, how do you deal with complex mixtures and multiple disease endpoints. You have to consider the interaction of the product with behavior, addictiveness, and appeal. Individual risk assessment may conflict with population harm, so people may

have a less risky product, but you may have more people using the product.

What effects are reversible? How do you deal with the population heterogeneity for tobacco use history? So is a 30-year smoker going to be different than a 2-year smoker? And their inherent disease risk, we keep doing studies on healthy people, but smokers are not healthy people. And then you have to think about the unintended consequences.

So I'll stop there and be ready for your questions at the discussion session.

(Applause.)

DR. DRESLER: Thank you. Talking about just a few challenges there.

Okay, so we're going to actually do the discussion now, so I know it said we were taking a break, but let's go ahead and have the discussion since we are still a bit ahead of schedule. So if I can have the speakers come on up, please, and sit up at the table. And then there's cards to write questions on, so if you want to write a question, raise your hand; you'll get a card and then share those in.

(Pause.)

DR. DRESLER: Okay. Oh, thank you. And Stephen Roberts, thank you. Thank you.

(Pause.)

DR. DRESLER: So I'm going to ask Dr. Roberts, if you just want to -- I know you're going to be speaking after this, but if you want to do like a quick introduction for yourself so people know in the room, and then Dr. Rice, I don't know if everybody was here yesterday, and if maybe you would just say something.

DR. ROBERTS: Okay, great. Yeah, Steve Roberts. I'm a toxicologist and risk assessor at the University of Florida. I direct the Center for Environmental and Human Toxicology there, and I've been engaged in risk assessment methodologies and risk assessment activities for 20 years or so.

DR. RICE: Glenn Rice. I'm a risk assessor for the U.S. Environmental Protection Agency.

DR. DRESLER: Okay. All right, so here's some questions. So let's start out with some of the advancements in risk assessment come from occupational and environmental health, so there are similarities and differences in those applications of risk assessments. So could you please just comment on these following aspects: the application of exposure response data to high end tobacco exposure scenarios, i.e., direct inhalation?

So what is the application of exposure response data to high end tobacco exposure scenarios? Does that make sense to

anyone, that you can -- yeah, that -- so who wants to jump in?

Vince, I see you going for the red button. Thank you.

DR. COGLIANO: Okay, thank you. I guess I'm going to -- you ask does the question make sense, so the way I understand that is that we have occupational and animal data at high-dose levels, and is that really applicable to the smoking scenarios that you have. I would actually say that's probably more directly applicable than what we do at EPA with environmental exposures because you are dealing with, you know, products that are intended to have a particular physiological effect on people and that they're supposed to be able to -- they're sensing that they're exposed to chemicals rather than at the trace levels we're often seeing in the environment, so I think you're actually probably closer to the range where some things were tested in animals or where there might be occupational exposures in the tobacco situation.

DR. DRESLER: Okay.

DR. SHIELDS: So it's certainly appealing. The big issue is do we understand how to assess exposure from cigarettes, so if you're going from the occupational world and saying, well, cigarette exposure is like this in that workplace, it's based on a smoking machine, and there's a lot of problems with that, I mean, especially if it's the FTC method, the earlier methods.

And so when we say, oh, the yield of the cigarette is X and in the workplace they get exposed to Y, we're really not comparing similar things. And then, of course, there's the route of exposure and the target organ. So there's some appeal there, but there's also a lot of limitations; that's my negative side. As you can see, I'm pretty negative about things.

DR. GENTRY: I guess I'd just further comment on that. I think a lot of the things you outlined we struggle with in the environmental situation as well. We don't have data to cover every individual situation, and certainly, in the environmental situation, we're charged with being protective of the population. So the risk assessment paradigm I talked about has the flexibility to talk about, in the risk characterization, uncertainty, variability, we've talked about different ways of quantifying that over the last day. Sometimes you don't have the data to quantify that.

Even in the environmental situation, we have the opportunity to talk about, in a qualitative way, what are the assumptions, where have we made assumptions, where do we not have data, how could that impact the conclusions and recommendations we're making, and I think that's something when Vince talked about the risk manager and making those decisions

that are protective of public health. All of those things have to be considered whether we have data or not.

DR. DRESLER: Okay. Dr. Gentry, what may be strengths and limitations of using a daily inhalation rate to estimate exposures to inhaled tobacco product constituents given that smoking is an intermittent exposure series of events? What other available approaches may be applicable? So let me read that again because just to -- please, I get to read your handwriting, so if you get -- cross out words and I have, you know, those carets that go up like that, it's hard to follow, so get another card if you want to, so please.

Okay, what may be strengths and limitations of using a daily inhalation rate to exposures to inhaled tobacco -- I'm sorry. Let's start again. What may be strengths and limitations of using a daily inhalation rate to estimate exposures to inhaled tobacco product constituents given that smoking is an intermittent exposure series of events? What other available apparatuses or approaches may be applicable?

DR. GENTRY: I think what I showed is probably the most simplistic approach. I mean, we want to compare these levels of constituents to these toxicity factors if you put it into the risk assessment paradigm, and so they've got to be expressed. Our toxicity factors are developed expressed as a

volume of air, so certainly the inhalation rate is a starting point to characterize what you might be exposed to. I think we've had discussions over the last few days of, you know, that overall continuous exposure metric may not be appropriate. We've certainly had people talk about tools that could be applied within that paradigm to refine what that exposure looks like considering peak exposure and changes. We've talked about pharmacokinetic models, CFD models, so we certainly have tools that could be used to refine that simplistic approach.

DR. DRESLER: Okay, I have a ton of questions, so thank you.

So, Dr. Shields, if you regulate ventilation, how do you keep people from altering the products themselves?

DR. SHIELDS: Well, sure, some people could do that. They could take a pin or something like that and try to make the smoke less harsh. That's something that we can -- would be an unintended consequence, and that's part of the research agenda to figure out whether people figure that out and start to do it.

DR. DRESLER: It might be hard to do that as sophisticated as those ventilation holes are.

DR. SHIELDS: That's exactly right, so -- you know, but you could get -- I mean, I guess you could. I never tried it,

but that would be one of the things that would have to be tested before -- I would think before regulation would go in.

DR. DRESLER: Okay, all right.

Is the focus on risk of cancer underestimating risk for other health outcomes? So, I mean, this is likely because cancer exposures seem to be much longer term, so is the risk of focusing on cancer underestimating the risks for other health outcomes?

DR. SHIELDS: Yes. You know, I think that all of these things, when I talked about assessing all morbidity and mortality, there could be differences in risks to smoking mothers, you know, or pregnant women versus kids versus adults and respiratory disease and cardiovascular disease and all of these. That would be part of the consideration, even when you don't have the data, and so maybe you can't get very far but, you know, all those diseases are important.

DR. COGLIANO: I think I would say yes as well. At EPA we did an assessment of lung cancer and other respiratory disorders from passive smoking in the early 1990s, and I think we found other diseases were actually higher in incidence, and I think there are a lot of studies that show cardiovascular diseases, you know, higher than the total cancer burden from cigarettes.

DR. SHIELDS: The example that I can think of is, you know, electronic cigarettes, so there are a lot of people who will say that the lung cancer risk from electronic cigarettes should be a lot lower than cigarettes, but I've also heard people say, and I won't tell you whether I agree or disagree, that their concerns for cardiovascular and pulmonary toxicity are as high as the cigarette smoke. I think that's based on no data either way, but that's an example of if we focus just on cancer, we could miss the boat.

DR. GENTRY: The paradigm I talked about for the environmental situation doesn't just focus on cancer. I gave that example, but we look at cancer and noncancer endpoints and look at the most sensitive endpoint for either of those categories. So you're developing a paradigm that if you're focusing on the most sensitive endpoint for that exposure pathway, it should be protective for the other endpoints that you're concerned about, and then looking at both noncancer and cancer together, not just one.

DR. DRESLER: Just to follow up, would you do that across products then? Because I'm thinking now in the regulated environment is water pipe, cigars, cigarettes, and electronic cigarettes. So now put that mixture into what you're going to regulate and how you decide which health risk.

DR. GENTRY: Well, but I think with most of the pathways to submission, it's a comparable product, right? When you're doing the substantial equivalent submissions, you're looking at a comparable product to what your new product is. So I think, in those comparisons, you're not trying to estimate the actual risk; you're trying to look at the differences in risk across two comparable products.

DR. DRESLER: And so I think for SEs, for substantial -- but perhaps not for the other pathways, so it might not -- okay.

So many of the OECD test guidelines, as written for hazard ID, many of the OECD test guidelines are written for hazard identification, and that came up some yesterday. For risk assessment, mode of action is often important and should be addressed by experimental designs that are different than the OECD guidelines. How does this impact judgments of studies that include assessment of compliance with OECD guidelines when one is doing a systematic or weight of evidence review?

So the question is how does the OECD guidelines impact judgments of studies that include an assessment of compliance with OECD guidelines when one is doing a systematic or weight of evidence review?

DR. GENTRY: When you're conducting a systematic review

and you're trying to evaluate the study quality, certainly for those types of studies where you have an OECD guideline, that's kind of the gold standard that you compare to for quality. There are other guidelines and approaches. OHAT, within NIEHS, has been working on systematic review approaches, and there are proposed methods for evaluating quality outside of those types of studies that have a guideline.

So some of the things that are part of the guideline comparison certainly could be extended to other study types, but there are some guidelines in the published literature and with other regulatory agencies for how you address some of those studies that don't necessarily have a guideline.

DR. DRESLER: Okay.

Dr. Cogliano, the EPA has experience with -- where did I go? With comprehensive -- or does a comprehensive systematic review approach capture more evidence than other approaches? So what is the difference in the level of effort of different approaches compared to a systematic review, and does it capture more evidence than other approaches?

DR. COGLIANO: I think it does capture some more studies. I don't know that they're necessarily the critical studies or that our older assessments really were deficient and missing a huge study, but part of that was we also have peer review and

public comment that provides opportunities to make sure we're looking at all of the pivotal or the high-quality studies.

Now, I think I would say that systematic review is taking more of an effort, more resources. I'm hopeful that that will change as we get more experience with it, as we get better tools, as we get perhaps more automation and are able to, you know, screen a large number of studies and sort them into high and low quality and have that all recorded electronically and our people get more experience with that. Right now it is higher in effort, but hopefully that will decrease.

DR. DRESLER: Okay. What types of mechanistic models would be appropriate for hazard characterization and tobacco risk assessment? What mechanistic models would be appropriate? I'm not sure who that would be for.

(Off microphone comments.)

DR. SHIELDS: Well, let me just say that I think that -- yeah. We may be touching on that in the next session a little bit in --

DR. DRESLER: Okay.

DR. SHIELDS: -- George's talk and my talk about -- we're focusing on risk, but I think part of that is going to be some hazard characterization and some potential approaches to doing that, so maybe we'll save that.

DR. DRESLER: We'll get -- okay.

DR. SHIELDS: Do that for the next one.

DR. DRESLER: I'll save that for the next one, so --

DR. SHIELDS: And if we don't address it, then they can ask it again at the panel discussion after that one.

DR. DRESLER: Okay. What experimental tests are available for testing addictive potential in the laboratory for comparing one cigarette to another cigarette?

DR. SHIELDS: So as part of a current multi-programmatic grant we have now at the University of Minnesota, in Virginia there's a guy named Warren -- I'm sorry. It's not Warren, it's Andy Harris, where they actually have abuse liability models that they validated for the addictiveness. So, for example, they can inject extract -- they did this with the smokeless tobacco and published this -- and see the behavior of the animal trying to get more or less of the product, and then they also have -- I know this sounds awful -- they could insert directly into the brain various extracts and look at the behavior. So there's behavioral economic models for animals that are actually pretty well validated and now being used for tobacco.

DR. DRESLER: That research has been going on for quite a few years --

DR. SHIELDS: Yeah.

DR. DRESLER: -- of looking at how do you do that.

DR. SHIELDS: Yeah. And we're funded by NCI to do that.

DR. DRESLER: Okay. Dr. Gentry, is there sufficient data on estimated risk for the constituents in the abbreviated HPHC list for cigarette smoke to estimate the relative contribution to risk for a conventional cigarette? So did that make sense? Do you want me to read that --

DR. GENTRY: I think what they're asking is I mentioned the decrease for portable lists within the 93 HPHCs, and I think for the majority of those compounds, those probably represent the chemicals where we have the most data.

There's still some challenges with some of those I mentioned, you know, maybe not having the appropriate data by the right route of exposure and so how do we handle that; do we do route-to-route extrapolation, but I think for the majority of those compounds, we have the adequate data to look at toxicity.

DR. DRESLER: So what do you do for, say, electronic cigarettes or the flavors that are being inhaled and there are no reference inhalation values for them?

DR. GENTRY: Well, I think we have to understand, a lot of those fragrances are grass; they're generally recognized as

safe for the oral exposure pathway, but I think we have to understand the potential consequences when we move from oral intake to inhalation. Maybe we can look at in vitro type studies or shorter-term studies to get a better idea what the differences may be between route to route. Some of the modeling may be applied to understand that, and certainly the TTC looks at some of that with the structure of the compound, you know, are there certain components of that structure that we would be concerned about from a toxicity standpoint.

DR. DRESLER: So Peter?

DR. SHIELDS: Yeah, that's right, but you also have the added thing about the heating of those compounds in the grass, so what's happening in the lung, and is there byproducts of that that I would just add to that, which is potentially an important point.

DR. DRESLER: Okay. What approaches may be used to consider whether the hazards identified capture the total risk of exposure to a complex tobacco mixture? So what approaches could be used to consider whether the hazards identified capture the total risk of exposure to a complex tobacco mixture?

DR. SHIELDS: Again, I'll try and touch on that a little bit from the point of view of cancer risk in a talk coming up,

but it is challenging and there's a -- there are lots and lots of chemicals, and many of which we don't have any information on so, you know, how can we get some information to address those is going to be an important issue.

DR. RICE: I think not only that, but I think you also have the concern that some of the models that we actually use for component mixture, you know, component assessments of a mixture, are really relatively simple. We often don't have interaction data that you'd really want to look at, so I think those types of efforts would be very difficult. And if you have the whole mixture information, I would submit that you would probably want to use that.

DR. COGLIANO: I think if you had perfect information, you might also find there's a lot of susceptibilities that take the risks that you do know but say that they're much higher in some subgroups of the population than in others, and we're missing that probably, you know, not any good way to know that until you actually do the right study.

You know, there's emerging methods for handling big data and very complex questions. I think the issue is the lack of data to be able to put into those models, you know, so you've got from cigarette smoking, benzene leukemia, aromatic amines, and bladder cancer and -- you know, and CO and heart disease

and benzopyrene and lung cancer, and I don't think we have the data, really, to isolate out different chemicals with different disease risks even if we had the models, and so maybe that just can't be done as part of this process, but it would be important if it could.

DR. DRESLER: Maybe not yet.

Please discuss possible approaches that may be used to evaluate local cancer effects for the oral mucosa associated with the use of tobacco products. So possible approaches that may be used to evaluate local cancer effects for the oral mucosa associated with the use of tobacco products.

DR. SHIELDS: So there's a number of different established, as well as emerging biomarkers on oral mucosa, and I would just add that there's a lot of people interested in nasal epithelial as being representative of what's going on in the lung because there's emerging data showing that you can just, you know, do nasal swabs or lavages, and that's pretty representative, but you know, there's studies for smokeless tobacco going on, you know, in terms of gene changes, epigenetic changes, changes to the microbiome, so that's a pretty active area of research.

DR. DRESLER: What approaches may be used to determine -- which chemicals? Which -- let me work on that one a little bit

more.

How may you compare risk under ISO and Canadian smoke conditions? So how do you recommend comparing the risks between those two?

DR. GENTRY: I guess not so much comparison, but those two regimes are meant to characterize bounds, so I mean, you could look at those similarly to the way in the environmental situation we look at mean concentration versus an upper bound. So that would be the comparison I would think of as helping, in some ways, to characterize all of the variability we've talked about in the last day and a half.

DR. SHIELDS: I mean, the problem is if you use any one smoking machine regimen, you may be using a regimen that doesn't represent exposures to anybody. And so we published a paper in 2009 looking at smoking machines versus behavior, and we weren't the first ones to suggest this, but if you wanted to try to get a little bit closer, as imperfect as it might be to understanding risk, you should be testing these products under multiple smoking conditions based on observational studies of how people actually use the product. So, you know, so health -- the Health Canada method is 100% hole blocking. Well, people don't smoke 100% hole blocking, but if you figure that some people smoke cigarettes with some hole blocking, some

with none, some take bigger puffs, some take lower puffs, you could potentially model three, four different smoking machine regimens to sort of represent the range of exposures and use that information.

Still lots of limitations, but I think that's substantially more appealing, even though it's not yet validated, than using a regimen to compare one product to another. I just don't -- I don't see how this concept of, well, let's use one regimen, and that gives us information, that's worthwhile? To understand the risk of one product versus the other, I totally reject that concept.

DR. DRESLER: Would you want to -- and anyone else want to address we also have electronic cigarettes or water pipes or cigars and ISO and high -- the Canadian. Anyone want to address that? Or just to say there isn't the ISO or Canadian Intense for these other products? So when you start looking at those -- but yeah, so we're looking across tobacco products, not just the cigarettes, right?

Okay, so what approaches may be used to determine whether -- I still can't do that one, I'm sorry. I think it's chemical. Let's see if it works. What approaches may be used to determine whether chemicals may be risk drivers in the exposure to a complex tobacco product mixture?

Does that make sense? What approaches may be used to determine whether chemicals may be risk drivers in the exposure to a complex tobacco --

DR. RICE: Yeah. Go ahead.

DR. COGLIANO: Well, I think it's similar to what we do when we're looking at like mixtures at a Superfund site, you might have dozens of chemicals. You measure the exposures, you look at the reference dose or the flow factor, and you see which one actually is the -- which ones are the highest ones, so that if you reduce the exposure to those, you generally reduce the exposures to many others.

Now, I think in a Superfund site, I can see that it's, you know, it means you're removing 6 inches of soil, and that reduces everybody's exposure pretty much to all of those chemicals. I can see that if you're talking about two different tobacco products and you're decreasing some and increasing others, you might have two different risk drivers in two different products, which is a little different from the environmental scenario. But I think what, to me, what that means is just what are the two or three chemicals that are -- that you really need to focus on to reduce most of the risk.

DR. RICE: I was thinking of a similar analogy with drinking water disinfection byproducts. Essentially what we do

in the U.S. is we regulate certain disinfection byproducts, and those are created when we add oxidants like chlorine to drinking water, and it reacts with natural organic matter to produce these putatively harmful disinfection byproducts.

We have two relatively large classes that are regulated in U.S. drinking waters. The trihalomethanes and the haloacetic acids, so two groups that we regulate the levels of in drinking water with the assumption that we control some of the risks of some of those disinfection byproducts. So it's similar to Jim's example.

Now, there are tradeoffs when you do that, and I think that would be the complicated part of regulating different tobacco products. You wouldn't experience some of those different tradeoffs as you're removing some products or reducing their levels, others may increase.

DR. SHIELDS: So it's an important question because you can imagine product standards attempting to have a product that lowers delivery. It would be nice to know the driver so that's where you start with. But it becomes really complicated because it's, well, how do you determine that because it's a complex mixture that people get exposed to, so when I was showing the concept of using the EPA slope factors, it kind of turns on the head what we think controls the drivers for lung

cancer, which is usually TSNAs and the pHs, but that analysis, you know, flipped it around and said, you know, it could be metals, and then there's also the endpoints.

So factors in nitrosamine may drive lung cancer, but if aromatic amines are driving bladder cancer and you can get a product standard that changes the aromatic amine exposure, maybe that's the driver you want to go after. So it gets complicated as soon as you sort of open up the hood.

DR. RICE: Just to go a little bit further with that. I think we had a question earlier about different types of risks associated with these different products, and I think, to do an assessment, you know, a comparative assessment like that, you would really want to try to get as many of those risks on the table to make that comparison as possible, and your data are not going to be real clean. You might have data that you're very confident in for one effect, data that you're not very confident for another effect, but I still think it's really important to at least acknowledge to your risk managers that are making some of those decisions that, you know, these are the different health outcomes that are in play potentially.

DR. DRESLER: And I think we probably do know that pretty much in the U.S. for, say, the top four causes of death in the U.S. We pretty much do have those attributions to cigarette

smoking. So within tobacco control.

Okay, so one other one. I have two, the two questions I want to ask hopefully kind of quickly. So we were talking about similarities and differences of application of risk assessment methods, and in particular, I'd like for you to address to the unique exposure pathways, like dermal application of dissolvable products or vapor versus particle inhalations. So what are similarities and differences on different risk assessment methods for when you're looking at unique exposure pathways?

DR. SHIELDS: Well, if you look at unique exposure pathways, there's two things you have to consider. You have to quantify the dose somehow for that pathway and be able to do that. And the other is that, you know, a basic tenet of toxicology is that toxic effects and the toxic potency can vary by route of exposure, so that has to be taken into consideration as well.

So as we look at other pathways besides inhalation, we have to be mindful that we have to now come up with toxicity values for those pathways that are relevant for those, and I think Jim mentioned or maybe it was -- route-to-route extrapolation.

I mean, there are some things you can try to do if you

don't have an experimentally derived toxicity value for that pathway, but it's an issue -- you have to take into consideration both, that route in terms of the exposure as well as the toxicity that might be invoked by exposure, that route as opposed to other routes.

DR. SHIELDS: So that's a real interesting question, so it reminds me of the example of the concept that we were testing is would there be some impact of shifting smokers to smokeless tobacco considering smokeless tobacco is a modified risk tobacco product compared to cigarette smoking? And if you look at oral cavity cancer, the risk of oral cavity cancer from cigarette smoke is much higher than smokeless tobacco. So that local effect is something that would be very different and not just looking at, you know, the inhalation component.

DR. DRESLER: So last question, Peter, for you. When you were talking about the ventilated cigarettes that were introduced in the U.S., when were they introduced into the U.S., and then when was the increase in adenocarcinoma?

DR. SHIELDS: So the use of filtered ventilation was a little bit in the marketplace in the 1960s. It became widely popular in the 1970s, and then over the next 20 years almost all cigarettes on the market were getting increasingly ventilated. The adenocarcinoma change happened around 1990 in

men.

DR. DRESLER: In men.

DR. SHIELDS: Yes, in men. And women always had sort of a higher adenocarcinoma risk, but that's because we assume because they always started with ventilated cigarettes, they started smoking later, and they gravitated toward the "light" cigarettes.

DR. DRESLER: Okay. Any other questions? We're over 6 minutes, okay. So now we'll take a break, a 15-minute break. It is 10:36, so let's come back at 10 of. Gives you a little bit under -- okay, 5 of. Enough time to go through that long coffee line upstairs, okay. So to go -- and I didn't start out this morning, the bathrooms are at the end of the hall to the right, coffee is upstairs. And I understand the line can be long.

(Off the record at 10:36 a.m.)

(On the record at 10:55 a.m.)

DR. DRESLER: We're seeing, looking at your watches, we are a few minutes ahead, so we're going to try and aim for around a 12-ish sort of lunchtime versus the later lunch and know that when you do that math, that means we might get out a little bit earlier today than our estimated ending time, which is a good thing all the way around, especially if we consider

all the fantastic talks we've heard so far and what are coming, so I'm excited.

Okay, so we're on to Session 5, Risk Assessment: Cancer and Noncancer Approaches, and our first speaker -- oh, I probably should have waited until he came to the room, huh?

(Laughter.)

DR. DRESLER: He just walked in, so very good, thank you. So our first speaker will be Dr. Stephen Roberts from the University of Florida. He was on our panel that you just heard from, and he will be speaking on Data and Approaches to Evaluate Cancer Risk from Tobacco Products.

(Pause.)

DR. ROBERTS: Well, it's really a pleasure to be here, and what I want to try and do in the next 20 minutes or so is really give kind of a 15- to 20,000 foot view of the landscape of potential data and approaches that can be used to evaluate cancer risk from tobacco product constituents.

And this makes it go forward -- no, it doesn't. Oh, there it goes. All right. Little delay, that's fine.

Okay, so we're going to be covering four topics. I'm going to talk about what tobacco-related products should we be concerned about in terms of cancer risk, an overview of what kind of approaches are available to estimate cancer risk from

individual tobacco product constituents, what are some reasonable expectations in terms of how these cancer risk estimates can be used, and then finally, some concluding remarks.

So the first, this is, you know, what is the universe of chemicals that we're interested in when we want to consider possible carcinogenic constituents of tobacco smoke. First of all, what is the universe of chemicals that are in tobacco smoke and tobacco products? I don't know whether this -- unfortunately, I wasn't here yesterday. I don't know if this got talked about very much, but there are a number of estimates in terms of how many constituents there are of potential interest.

Hoffman and Wynder, IARC publication 1986, estimated about 3,900. A more recent study in 2011 estimated greater than 5,000, and the second edition of *The Chemical Components of Tobacco and Tobacco Smoke* puts it at about 9,600. So whether it's 3,000, 5,000, 9,000, it's a lot of chemicals. So if we're going to evaluate the potential cancer risk, we have a lot of chemicals to consider.

How many of those are carcinogens? The answer is very short: we don't know. We really don't have information on the vast majority of those. So let's rephrase the question: How

many do we know are carcinogens? Or I'll expand that a little bit: How many do we know or suspect are carcinogens? There have been some reports in the literature, analyses that have tried to go through this list of thousands of chemicals and pick them out. Probably the most germane example to point out here is one done by the FDA and coming up with the HPHC list, and they published a list of chemicals found in tobacco smoke that are known or potential carcinogens as defined by IARC, EPA, NTP, and NIOSH.

And let me just clarify, I'm using the word "potential carcinogens" kind of generically to cover the nomenclature that each of these agencies use in terms of identifying chemicals that might cause cancer, which would include terminology like "reasonably anticipated to" or "possible carcinogens," "probable carcinogens," those kinds of things. So the EPA assembled that list, and by my count, there's 79. So there's 79 known or suspected carcinogens on that list. Does that mean that's all of the known or suspected carcinogens in tobacco smoke or tobacco-related products? Certainly not. There's a lot of chemicals that we don't have any information at all on that.

What about nicotine delivery devices? And this is a little bit newer problem in terms of trying to figure out

what's in the emissions from nicotine delivery devices. This is probably basic information to everybody in this room; you know, the primary ingredients are some kind of humectant, flavorings, and nicotine. There is, as was pointed out, I think, a little earlier today, it's limited information on what's in e-cigarette emissions; there's not very many studies. The results are kind of variable, which may be due perhaps to different -- testing of different products, the conditions under which the products are used, temperature, voltages, those kinds of thing, different test methods and so forth.

But at least the information that we have available today suggests that some of the substances on the HPHC list have been found in some emissions from e-cigarette devices, including some things that are known potential carcinogens, which include tobacco product-specific nitrosamines, NNN, NNK, formaldehyde, acetaldehyde. So there's reason to encompass e-cigarette emissions, as well, when we're beginning to think about trying to identify potential carcinogens in these devices.

Okay, so what are we going to do about all the other chemicals that are potentially present in tobacco-related products and trying to decide whether they might pose a risk of cancer, because we got a number of 79, but we know there potentially could be many more. What kind of tools might be

available to do that? One is the Tox21 approach, which is basically a high-put screening method using assays for biological events, and since we're talking about cancer, they would be key events related to carcinogenicity. So in order to do this, we would really need to know what are the biological chains of events that lead to cancer for the kinds of cancer that we're concerned about from smoking and the use of tobacco products.

So we would have to know what those are, we would have to -- then we could pick out key events, could do assays for those and try and determine whether or not an individual constituent produces the kind of effects that might plausibly lead to cancer. And the advantage to this is that you can screen chemicals very quickly and relatively inexpensively. The high-throughput methodology, you can screen literally thousands of chemicals at a time. So that's the good news.

The disadvantage is you really have to know what to look for in those assays, you have to be able to pick out assays that are going to be capable of detecting carcinogenicity really produced by conceivably a variety of different modes of action, that sort of thing. So unless you know that, unless you can produce a comprehensive list of in vitro assays that you can screen, you know, there's a pretty high likelihood that

you're going to get some false negatives, there's going to be things that you're just not going to pick up because you haven't done the right assay.

Quantitative structure-activity relationship, this is basically where we take computer models to predict carcinogenicity based on molecular descriptors. They rely on databases of chemicals with known carcinogenic effects and their structural properties, and then are used to try and predict what structural properties might produce cancer from other chemicals and unknowns. That's a potential way to try and look through some of these other chemicals and try and identify carcinogens.

And then there are other processes like chemical category and read-across, and read-across is where you look at, you're trying to find analogs or chemicals that are related to the one that you're trying to understand its toxicity, and it could be related structurally or through a number of physical/chemical properties, and if you have a bunch of analogs and you can compare them to the chemical, you can gain some insight about whether or not -- you can make some assumptions about whether or not this chemical is going to behave like the others and do some analysis in that regard. You know, read-across can work pretty well if you've got analogs, a number of analogs for the

chemical that you're interested in, and it works best if those analogs all are giving you a consistent picture about the toxicity. If you have similar chemicals and they have very different toxicity, then you're really not sure which is one that's representing the chemical that you're interested in.

So read-across, those kinds of things could conceivably be used if you've got analogs for -- without conflicting toxicity for some of the chemicals that you're trying to screen through this list for tobacco products. If you don't, QSAR might be another option to try and take a look, and at least try and pick out and identify the chemicals that you might be concerned about.

So once you've got your list of chemicals, how then are you going to estimate cancer risk? Okay. So I'm going to talk about four approaches here. First one is -- and I'll just apologize for the very first one. Exposure times cancer potency equals risk. This is really sort of the environmental approach that has been talked about this morning and described very well. I didn't want to label it as the environmental approach because I didn't want to sort of say this really only works for that, but this is the fundamental approach, and it was described very, very nicely by Robinan this morning. So it's used to write quantitative risk estimates typically used

for environmental contaminants.

There's margin-of-exposure approaches; I'll describe that. That is currently being used to assess risks for food additives.

Threshold of toxicological concern for food packaging materials and flavorings.

So these are all approaches that are being used. One that's, I'll just say is in development, is on the horizon, is the Risk21 approach, and this is of interest because it allows, in contrast to some of the others, in vitro data to be used for risk assessment, which is easier to obtain than in vivo data.

Okay, so let's talk about the first one, exposure times cancer potency equals risk, that cumbersome title, all right. For non-threshold carcinogens, you can derive a numerical cancer risk associated with a known or estimated exposure; again, this has been described very nicely this morning. You'd have to have a cancer potency estimate for that chemical and for that route of exposure, and it's usually expressed in units of reciprocal dose or reciprocal concentration in air for air exposure. And they're primarily derived from cancer study bioassays in animals is where we usually get most of this information.

There are a number of different ways that you can derive

those cancer potency estimates from cancer incidents versus dose data. Vince Cogliano described very clearly the approach that's currently being used by the EPA to derive a cancer potency estimate from cancer data where you fit the data and you do a linear extrapolation, that sort of thing.

I would just -- to that, I would just add that when you go to the sources that have been talked about where these data are available, things like EPA and so forth, there are some, I'll call them legacy cancer potency estimates have been derived using other methodologies like the linear, linearized multi-stage model, and so what you find is, because some of them have not been updated in a long time, as you pull these, you have to realize that you're working with cancer potency estimates that aren't all derived using a consistent methodology.

As I mentioned before, cancer potency factors for several chemicals are available from U.S. EPA IRIS database, CalEPA. The advantage of these are they're peer reviewed; they undergo quite a bit of scrutiny at input. There are also some values available from the literature, published literature, some other sources. Typically, they have less peer review.

But when you go about this process, you can go pull these numbers at least for some of the chemicals and on the HP -- I'm sorry, the high-potency potential -- I'm sorry, the HPHC list.

I believe there about 40 chemicals that you can get from this source, and Dr. Annette Santamaria is going to be giving a presentation, she'll probably clarify that, exactly how many you can get from these sources, and her presentation will be coming up later this morning.

You can calculate these cancer risks using deterministically or probabilistic methods, and I think I won't go into any detail because she's going to illustrate that, I think, probably in her talk, right? Yeah. Okay, good.

Okay, so what are the advantages and the disadvantages to this approach as we apply them to tobacco products? It gives you a numerical estimate of risk, you know, the risk of cancer from this under a given exposure scenario is one in a million, one in 100,000, one in hundreds, that sort of thing. So you get a numerical estimate of risk, so that's a quantitative estimate of risk, and sometimes that can be very useful. And there's also a regulatory precedent for the use of this approach. EPA uses this extensively; it's been used for years and years. So you certainly can point to them.

There are a number of limitations, and people who work with this methodology in the environmental field are very familiar with these because these have been discussed at length over the years. The cancer potency values are available only

for a limited number of tobacco product constituents, so we can do this with 40 or whatever the number is, and that leaves some unknown but probably very large number of carcinogens for which we have no potency information.

The potency estimations are -- estimates are derived primarily from animal studies given relatively high -- should be relatively high doses of the chemical, so the questions always arise: Are the animals good surrogates for humans? Are they giving us reliable estimates quantitatively? Are they giving us reliable information qualitatively in terms of the kinds of cancers that we would expect to see? They're given in relatively high doses, and this leads to the issue that's been discussed for decades is what is the shape of the dose-response curve at the low-dose range, and depending upon the shape of that dose-response curve, can have dramatic effects on the cancer estimates.

So we can do things like, you know, draw straight lines and that sort of thing, but that really is just a practical approach that's pursued in the absence of knowledge of what that shape really is. So that's always an uncertainty and a limitation.

The third bullet is that cancer potency estimates may be based on tumors that occur other than at sites where -- of most

interest for tobacco-related cancers. So the question can come up. If we got, you know -- typically what's done is you'll look at the animal data, and they may have two or three different cancers, and you'll look at the cancer that occurs at the lowest doses, and you'll model the potency based on that because when EPA uses this thing, they want to prevent all cancers, so they're okay with overestimating the risk of some cancers by using that potency factor, but they want to make sure they don't underestimate them. So the question is if you use a potency estimate for this and it's based on pancreatic cancer or something like this, you say -- and say, well, that's not really something I'm necessarily concerned about for the smoking, I really want an estimate based on a specific cancer that -- I'm interested in smoking.

And this really gets into the argument of how important is site concordance between the cancer that's produced in the animal and what you're trying to predict in humans. And there are arguments both ways. The argument, there's an argument that site concordance isn't necessary because there's really lots of reasons why animals -- a chemical can produce tumors at different sites in animals than it does in humans, so there's lots of reasons for that, and therefore, site concordance is very important.

There are some counterarguments by some folks that actually were most concerned about the cancers associated with the mode of action specific for these chemicals, and maybe the animal isn't a good model for that. So it's just an issue on uncertainty associated with doing that.

And the last one, and the big one that's been talked about, I think touched upon several times already, is if we're going to develop cancer estimates for individual components, how do we combine them? I mean, do we assume that there's no interaction, just simply add them, or are there potentially interactions that could result in, when you do them, that could result in a greater or lesser cancer risk when they're combined? So these are some of the limitations associated with using this approach.

Margin of exposure approach: That determines the difference or the margin between a known or estimated exposure and the exposure known to produce cancer in animals or humans. So you're going to derive a dose known to produce cancer; specifically, the terminology for that is some kind of point of departure, and you're going to compare that with actual or estimated exposure to see what the difference is.

And obviously, if there's -- if the exposure is much less than the dose known to produce cancer, that's a good thing. If

the dose is very close to the dose known to produce cancer, that's a bad thing. So the risk, in a sense, increases as the margin of exposure, the closeness of those two doses decreases, okay? And as I think was pointed out in Vince's talk a little bit earlier this morning, he illustrated that with a nice figure that showed that you can develop the point of departure from benchmark dose modeling, and you can do an extrapolation down to zero, you know, typically the cancer incidence. You might want to model or the response you might want to model is about a 10% cancer incidence, but this could be different in different situations.

And then you want to see how -- what the margin or the ratio between the actual exposure in that one is. And as some regulatory agencies have set rigid acceptable margins of exposure, we're going to be -- we're not going to be concerned, for example, if the margin of exposure is greater than 10,000 or 1,000, that sort of thing. It gives you a way to create some benchmarks for interpretation.

It provides an expression of risk, what I just said, that the risk decreases as the margin of exposure increases. It doesn't really provide a margin of exposure of 10,000 or 1,000; it doesn't give you a numerical probability of cancer, although you could derive one from a margin of exposure assuming that

the relationship between cancer and dose is linear. So if you assume it's linear, you could take a margin of exposure and come up with a cancer risk number.

It's been using -- been used for screening purposes in a number of different areas by setting a margin of exposure above which concern is minimal, and for example, as I just mentioned, EFSA considers chemicals with a margin of exposure greater than 10,000 to be low priority when they're looking at genotoxic and carcinogenic chemicals in food, okay. So it's kind of an approach that's being used for chemicals in food; maybe we can adapt that to tobacco constituents and make some decisions about which things are a concern or not a concern.

The advantages are it provides a quantitative assessment of cancer risk. There's regulatory precedent for using it, so that's somewhat reassuring. There are a number of limitations. The cancer potency values are available just as with -- these limitations are really the same limitations I talked about for the environmental approach, the exposure times cancer potency. You have to have a safe human dose to be able to do it for cancer and, you know, all the limitations that are associated with trying to derive those.

Another approach is the threshold of toxicological concern. This is a means to eliminate chemicals as being of

low toxicological concern if the exposure is sufficiently low. Basic process is you collect safe doses for as many chemicals as you can get, hundreds of chemicals. They can be safe doses based on a specific cancer risk, like 10 to the -6, for example, or it could be safe doses based on no effect levels observed.

You look at all of them, and you say, you start, you can rank them, and you can see at what dose which ones -- what dose is small enough that you don't exceed the safe level for any of these chemicals. You pick a dose near the bottom, and you establish that as your threshold. And what you say is, based on our experience when the dose is this low, we really don't see a problem with a chemical no matter what its effects are, no matter what its mode of action, any of those kinds of things.

Set the dose low enough, you can basically, with confidence, assume that chemicals that are exposed to below that will have a low probability of producing a toxic effect. That's the concept behind the threshold of toxicological concern. Obviously, the more chemicals you have data for, the better. This is an approach that's used by -- has been used by FDA for indirect food additives and by JECFA for flavoring, so again there's regulatory precedent for using this approach.

You can apply it to both carcinogenic and non-carcinogenic substances. And as I pointed out, you use thresholds supported by empirical data to show that the probability of a chemical, any chemical, having an effect if the exposure is below this threshold is low based on our experience with all of these other chemicals, okay. Now, you have to use the thresholds in conjunction with other information about the toxicity of the chemical. You can't just apply the threshold uniform. You have to -- if you know something about the chemical, you can't just say, well, if it's below this threshold, we're going to assume it's okay even though we know something about that chemical and there's actually potentially a problem below that threshold.

So that has to -- so generally, these are posed, are utilized in sort of a scheme or a framework where they're compared with what's known about a chemical if there is anything. For carcinogens, there's some tinkering with the threshold TTC approach; that's sometimes necessary because some of the tinkering is for known carcinogens. We're not going to eliminate them no matter what the exposure is. For others, if -- we're not going to eliminate them if we think they might be a genotoxic carcinogen based on some structural alerts. So that would be another reason to not include them in the

approach. Or another way to do it is for a genotoxic or something that might be a genotoxicity -- genotoxic carcinogen, we'll just use a lower threshold. So we'll set the threshold lower for genotoxic carcinogens. It's another way to deal with that issue.

Okay, the advantages that allow a risk decision to be made with little or no data on the chemical, it can be done fairly quickly, easily, precedent for regulatory use. It's best used because all the thresholds are typically in the microgram or sub-microgram per day range, the ones that are used for regulatory purposes, best suited for things that have low exposures.

It has -- and it just provides you a yes or no answer. It doesn't really give you a risk estimate. It just -- it's either okay or maybe not okay. And the other thing is you'd have to have a good estimate of dose, and there's been a lot of discussion, I think -- unfortunately, I wasn't here yesterday, but there was a lot of discussion about how we derive dose estimates for tobacco constituents, but you'd have to have a good estimate of the dose.

Risk21 is the fourth one I want to talk about. I'll just talk about this briefly. This is a figure that's taken from the ILSI HESI site. This is an approach to be able to use a

variety of types of information on toxicity and exposure to estimate risk, begins with a problem formulation step. Here it begins with -- it includes exposure information, and that can be of varying degrees of sophistication. It can include toxicity information, which is not going very easily here, which could be as limited as QSAR or TTC information, up to in vivo/in vitro or even including modes of action comparing these and trying to come up with an estimate of whether or not the exposure is of concern or not. The reason I include this is because in contrast to the other methods that I talked about, this method can use in vitro data, so that's -- and in vitro data can be obtained much more easily than in vivo data, so be thinking about trying to address all of these chemicals in tobacco or related products; it's an opportunity.

It's basically a margin of exposure approach; it can use in vitro data. It also uses exposure estimates relative to scenarios of interest or the risk question that you want to address, so that's very important to have realistic exposure information.

And so ideally, as we talked about a little earlier with the in vitro test, if you had an in vitro test and you had key events and you could identify concentrations, limits under which those key events do not occur, then you could come up

with safe levels of exposure for that constituent based on in vitro data, in theory, that would prevent carcinogenesis. One of the issues, though, is you have to then convert a concentration in an assay or an in vitro tissue into an exposure level for a human, and that can be a little bit tricky.

There is a lot of effort going on to try and be able to do those kinds of extrapolations to get, to convert a concentration, for example, in a tissue culture to a corresponding human exposure. It can be -- it's called reverse toxicokinetics. It's most easily done if you've got good kinetic information on the chemical, ideally enough from which you could create a physiologically based toxicokinetic model. But oftentimes, I think, probably the best -- maybe all you're going to be able to get for some of these chemicals, you're lucky if you get a half-life, and so that makes these reverse calculations a little tricky.

Advantage is that you can do assessment with little or no in vivo data. The limitations, as I talked about before, is really, I don't think we're ready, my personal opinion is we're not ready yet to understand all the key events that are related to carcinogenesis, and then there's also issues with reverse kinetics.

How are we going to use these risk estimates? I think, ideally, we would like to be able to have risk estimates that are accurate enough that we could use them to do the kinds of things that Peter was talking about a little earlier.

If we remove this constituent, we know what the cancer risk from smoking is, and we can -- if we remove this constituent, we could predict how much of a reduction in cancers we would see by doing that, that would be great. And there's also some advantages to doing some ground-truthing in terms of whenever we do these risk assessments, it's nice to know that the estimates we're coming up with have some basis in reality in terms of prediction of what's going on in humans. But there's lots of reasons why these may not be accurate, and that's that they require accurate information on exposures and doses that we may not have. The cancer potency estimates, again, are based on animal studies using high doses; there's the whole issue of the accuracy of extrapolation to low concentrations.

We have to recognize these ways that are developed are conservative. They're intended to be sure that they don't underestimate cancer risk, so that confounds sometimes a comparison with cancers that we're seeing in populations. And then, of course, the other issue is that we can't -- until we

can address the potential carcinogenicity of all the risk drivers, it's going to be difficult to have accurate risk estimates.

Okay, next slide. Best use for foreseeable future is probably going to be prioritization of chemicals, trying to identify what the risk drivers are and the highest priorities for more research and perhaps some management strategies. And then, of course, also of interest to this group is for risk comparisons among different tobacco products and the importance of being able to do that for regulatory purposes.

So last concluding remarks: The issue we're facing with these tobacco constituents is just the classical modern toxicology problem where we have lots of chemicals that we don't have much toxicity information for. What are we going to do about it? It's just not going to be feasible to do animal studies to get information on all of these chemicals; we're going to have to look for alternative strategies, and just as we're doing in other areas of environmental toxicology and so forth, we have to look for other approaches.

TTC, QSAR, chemical category and read-across, we have to really take a look at all of these approaches as ways that we can try and get some cancer risk information and try and do a better job at more comprehensively addressing the constituents

in tobacco products.

And that is it.

(Applause.)

DR. DRESLER: Thank you. As you were saying that, I'm thinking yes, and we need to be applying everything that you're saying we need to do in the future, we need to do it today, so how difficult this task is.

Okay, so our next presenter is Dr. George Gray from George Washington University, and he'll be speaking on Approaches to Non-Cancer Risk Assessment of Tobacco Products.

DR. GRAY: Good morning, everyone. I appreciate the opportunity to be here, and just like Steve, what I want to do is give an overview of the way that we think about noncancer risk assessment and look at sort of the pros and cons of different approaches that are out there and to look, as he did, at some of the challenges that we have because of a lack of information on a lot of our chemicals.

So what I want to do is talk about the challenge, what's out there now, how people think about noncancer risk assessment, and then talk about why we need to think about other ways to do things and be a little more willing to, I think, recognize uncertainty but to move forward in the way that we want to characterize chemicals.

So that list on the left is not meant to be read; it's a whole bunch of the chemicals that have already been identified as potential chemicals of concern. There's lots and lots of chemicals in tobacco products, and the important thing, the only thing that makes toxicology particularly interesting is that things vary in their toxicity, and understanding why that is and how much that is, is actually really pretty important. So we know that all of these chemicals are going to be different, so we can't treat them the same. We need to know something about them. We want to characterize them, how the risk that they pose.

And characterizing those risks is going to be important for a variety of things, including setting priorities, so this idea, which things are we going to spend more time researching and which ones are we ready to say something about now, which ones are the ones that we need to pay regulatory attention to now, and which are the -- how are we going to communicate to people about those risks, because some of the things that may be done to address tobacco risk could be communication interventions, too.

The big problem we have, and it's been alluded to many times, is that these chemicals vary in the availability of data, and not just of data, but of tox values. In my research,

in my group, we call them human health reference values, and these are things like acceptable daily intakes, reference doses, cancer slope factors. We look to authoritative bodies to provide those for us. And those processes are slow. We do 5 or 10 evaluations a year in EPA's IRIS program. Other programs around the world do them at a similar pace. Making up this list is -- got to move faster than that.

One thing I do want to take a minute to do is remind us why we even -- why do we split cancer and noncancer both in the way that we're talking about it this morning and in the way that risk assessors have thought about this for a long time?

And what I want to remind us is the fundamental idea is there's a difference in underlying biology between the processes that cause cancer and the processes that lead to other effects. And the big idea is that cancer effects are still stochastic; they have a random process to them. They're the response of individual events in a single cell. Those events become auto- or self-amplifying, they're not dependent on dose.

When we think about noncancer effects, the severity depends on the dose. The more exposure you have, the more severe the endpoints. With cancer, the idea is if you have more and more exposure to benzene, your leukemia doesn't get

worse and worse; the likelihood of getting leukemia goes up. So there are fundamental underlying biological differences. So in noncancer, what we're looking for is trying to identify, in most cases, levels of exposure where these bad things won't happen.

So I'm going to talk about three kind of approaches that have been used. Some of them are the same ones that Steve just talked about on the cancer side. I want to talk about what I'll call the uncertainty factor approach, and this is what the Food and Drug Administration uses to setting acceptable daily intakes. U.S. EPA calls them reference doses. The Agency for Toxic Substances and Disease Registry calls them minimal effect levels.

(Off microphone comments.)

DR. GRAY: What's that?

UNIDENTIFIED SPEAKER: Risk --

DR. GRAY: Risk, MRLs. And other organizations around the world have similar sorts of names for these. I want to talk about the margin of exposure and then the idea of continuous functions for noncancer effects.

All of these approaches somehow begin with something that is commonly called a point of departure, and the point of departure is something on the dose-response curve of what's

called the critical effect in the critical study. We'll talk about those in a minute. But the critical effect, in most cases, is the most sensitive effect that's been found in some sort of toxicological studies, and by definition, the critical effect happens in the critical study, the critical study being a particular analysis, a 2-year bioassay of the chemical in mice and rats.

There are two primary ways that we see point of departure being set on the noncancer side, either using the traditional no observed effect level or no observed adverse effect level, or using the benchmark dose approach.

So these are just these graphs to show us the no observed and the lowest observed adverse effect level. An important thing about those is those have to be specific data points in a study, so they are specific places where measurements have been taken and endpoints evaluated in an animal study.

On the right is the benchmark dose approach which we've talked about, which uses continuous models on that same data, but to estimate a consistent starting point, often something like a benchmark dose for a 10% response or the lower statistical bound on that, the BMDL.

So the uncertainty factor approach simply takes that point of departure that comes from -- we take all of the data that

have been found on a particular chemical, sort them into the different endpoints that have been caused, the doses at which those were seen, either no observed effect levels or benchmark doses, and usually the most sensitive one is chosen, and the point of departure for that is then divided by a number of uncertainty factors.

This approach grew out of actually -- one of the first publications focusing on this came out of the Food and Drug Administration in 1954. So this approach has sort of built up over time and been a pretty traditional approach in regulatory agencies around the world. And it takes that point of departure and divides it by uncertainty factors that can take out a value of somewhere between 1 and 10, and the specific uncertainty factors that are used depend on the attributes of the study that you're starting with.

So when this is done, you come up with an estimate of what FDA calls acceptable daily intake. Again, EPA calls it a reference dose; others call it tolerable intakes or minimum risk levels. FDA's definition of what you've done in this process where you've take that point of departure and then divided it by these factors: "An ADI is the amount of a substance that is considered safe to consume every day over the course of a person's lifetime." EPA's definition of a

reference dose is quite similar. It doesn't use the word "safe"; it uses "likely to be without an appreciable risk of deleterious effects during a lifetime." But in both cases, the idea is that this particular approach helps to identify a level of exposure that people would be exposed to every single day for a lifetime without adverse effects occurring.

These approaches can also be rolled up in a way, and there have been attempts to do this in the tobacco literature to take constituents of tobacco or constituents of tobacco smoke and calculate for each one of them what's called a hazard quotient. The hazard quotient is a simple ratio of how much exposure there is, and this would be a daily exposure, to your reference value, whether it's acceptable daily intake or the reference dose.

And the idea is, is that a hazard quotient that's much less than 1 means that your exposure is much less than your reference level, and everything should be okay. When things get a little closer to 1, the implications in the interpretation becomes a little trickier. Something else that can be done to roll these up is to calculate something called a hazard index, and a hazard index simply adds up the hazard quotients from all of the combinations of a mixture. And this is some of the things that Glenn Rice talked about yesterday.

This is one way to look at mixtures.

And again, this has been done for cigarettes in a couple of papers that are out there. So that's the approach of, sort of the traditional approach of taking the dose that did nothing or very minimal effect in the most sensitive animals that have been tested for the most sensitive endpoints and then dividing it by these uncertainty factors to provide a margin of safety of some sort.

Another approach is the margin of exposure, and in the case of noncancer effects, it's calculated exactly the same way as Steve talked about; the point of departure is divided by the exposure. So the point of departure usually comes from tox data. It could come from epidemiologic data, if you had it. It could even come from in vitro data if you had a way to convince yourself that you could do a good job of calculating that. But again, the point of departure is either no observed adverse effect level or benchmark dose or bound on that benchmark dose.

So margins of exposure, this is the way Health Canada defines it. It's "the magnitude of the ratio between the level at which the critical effect is observed," so that's your point of departure, and the human exposure to the substance. And here, in contrast -- remember, a hazard quotient, we like

numbers that were much less than 1. In this case, we like the bigger the number, the better. So hazard quotients will be in -- excuse me.

Margins of exposure will be greater than 1. They will be in the neighborhood -- they might be in the neighborhood of tens, hundreds, or even thousands, the ratio between the human exposure and the point of departure from the animal studies.

The European Food Safety Authority uses this: "The margin of exposure is a ratio of two factors which assess for a given population the dose at which a small but measurable adverse effect is first observed," that's our point of departure, "and the level of exposure for the substance considered."

And margins of exposure, in some ways, as I look at risk assessment, I'm starting to see these used more and more around the world. So you see I've got some organizations at the bottom. The European Chemicals Agency, the Norwegian Scientific Committee, the Australian Department of Health and Aging, parts of EPA use the margins of exposure, the pesticide office most prominently. So this is an approach that is, in my view, being seen more and more. Another approach to doing this is rather than coming up with levels of exposure below which we sort of think bad things won't happen is to recognize that even noncancer effects, for a variety of reasons, either because of

the way that they cause their adverse effect, the biological steps that are going on, or simply because of a wide range in variability in the human population, may not actually have identifiable thresholds, and the thing that we may want to do is to begin to use continuous dose-response models much like are used in the cancer world for noncancer as well.

So one example of this I point to, this was something that was identified in a National Academy of Sciences report called *Science and Decisions*. And this other -- the other is a probabilistic framework for dose-response assessment of human health effects; there's published environmental health perspectives. The International Programme on Chemical Safety has looked at this. There's a variety of groups who have suggested that maybe what we want to do is to really get -- bring cancer and noncancer risk assessment closer together so that in both cases we can make quantitative estimates of the likelihood of an adverse outcome.

So now I do the fun part where I say what's good and what's bad about each of these. You know, Steve did this, and no approach is ideal for every setting that we want to use. So the traditional uncertainty factor approach, what I call here the ADI or RfD approach, has been used for a long time, and people are familiar with it. As far as we know, it has not led

us terribly astray; it's hard to find solid examples of places where when exposures were well below those levels bad things have happened. It's also relatively -- it's got some idea of how we apply it to mixtures, adding up those hazard quotients to come up with a hazard index; it's been used widely.

Now, I personally don't like that approach. It reminds me of like doing a world tour and coming back with a pocket full of change and adding up shekels and euros and yen and say, oh, look, I have 22 money. It's not an interpretable -- the hazard index thing is not interpretable as we've done it across chemicals. That's my personal view. But it's also easy to communicate. Are you above or below a level that has been said cannot -- a level that should be safe. And remember, FDA actually uses the word "safe" in their definition of an ADI. When your exposure's below there, that's good. Your exposure's above there, maybe we need to do something. So there are limitations here. Like all of these, it requires data.

And another limitation that I would put there is that as long as we're going to look to authoritative bodies to provide these for us, they're also not going to -- even when there are data, there are not going to be nearly as many of these as we'd like to have. There are a lot of science policy judgments in their development. I'm going to show you an example of that in

just a second.

Most toxicologists and risk assessors are pretty comfortable saying if your exposure is well below the acceptable daily intake, things are probably okay. If you come to them and say my exposure is exactly at the acceptable daily intake, they're not so sure what to tell you. Don't know if you should run around like your hair's on fire or just relax there, too. There's some amount of safety built in. Interpretation of these numbers: These aren't risk numbers; these are simply safety assessments with the notion being if you're well below, you're okay. Once you get nearer these numbers, it's harder to interpret. So it doesn't provide us any kind of an estimate of risk, and another thing that I want to show you is one of the concerns here is that it doesn't allow you to judge, in many cases, the benefit of an intervention that you might take to reduce exposure.

So just let me show you two of these things. The science policy judgments, so these are noncancer reference values from five different -- excuse me, four different organizations for the same chemical, and these were done at approximately the same time, so they had the same data. Each group has the same data available to them.

They're doing that approach we said where you identify

point of departure, you divide it by uncertainty factors, and you report out what you think is -- and I'm going to use the term; they won't -- but a safe level of exposure. And what I want you to see here is that for these three different chemicals, they don't come to the same conclusion. And that's not because any one of them, as I said, has data the other doesn't have. It's not because any one of them is smarter or cares more or cares less about people. It's just that there are judgments that have to be made in this process. Different organizations make different judgments.

So here, I just want you to see that even when we have those numbers, some of them you can see there are tenfold or larger differences between these organizations. And there's not even a really strong pattern of any one of them being either more conservative, that is having lower levels, or less conservative and having higher levels. The point here is that the development of these particular values has additional things besides science that go into making those judgments.

And now I want to show you what I meant by this inability to use these kinds of approaches to look at interventions. So let's say that the reference dose really does identify a safe level of exposure; no, it doesn't. So there's our reference dose. And so the idea is, is below that, our exposures to the

left of that red line are low risk or even no risk, right? And above, we've got concern about them. Now, imagine that we've got a level of exposure that's up there; it's a high risk exposure, and we want to reduce that. If I reduce it from there to there, there's no benefit because I'm still above that reference level. Similarly, if I've got an exposure that's down below, and I can reduce it, there's no value to it. So we can't use this to say comparatively how much benefit am I getting from reducing exposures if they don't cross that RfD line, and even if they do, we can't make any quantitative estimate of how much risk we've avoided, how many injuries have been prevented, how much morbidity has been avoided.

So let's look at the MOE approach. So again, the MOE approach has a history of use, and as I said, I just -- as I look around, I'm seeing it used more and more around the world. It seems to be something -- in some ways it's risk assessors saying I don't want to get -- you know, I don't want to get blamed for choosing which uncertainty factors are right or which ones are wrong. I'm just going to tell you where the science ends, which is my estimate of a point of departure, and then you people who have decisions to make, you decide which margin of exposure is appropriate. So it's got a history of use, people are pretty familiar with it, and again, it's

relatively easy to communicate. You say your exposure is a thousand times lower than our point of departure, and we consider that to be a pretty big range. Now, again, this has data requirements. We still need some kind of data in order to come up with a point of departure. Again, it doesn't provide us an estimate of risk. All it provides us is a ratio between the level that did nothing or little in animals and what people are exposed to.

It doesn't have strong theory for how we would use it in mixtures, how we would combine that; it doesn't even have sort of the -- in some ways the underlying support of the hazard index approach, and there's really no scientific guidance on how big a margin of exposure should be.

So here's some -- you know, the question of how big does a margin of exposure have to be. You know, one of the first things we do is just go to these simple orders of magnitude, you know, a hundred, a thousand, you know, it could be 233 is the right one. There's no normative basis for choosing the right size.

Here's a quote that came out of -- the Consumer Products Safety Commission had an outside group look at phthalates in children's toys. They used the margin of exposure approach, and they said usually margins of exposure exceeding 100 to

1,000 are considered adequate for protecting public health. So their idea was 100 to 1,000. In ecological risk assessments, EPA uses 100 to 1,000 as kind of an appropriate level, but again, that's a pretty wide range. And something we can't do with this, again, in judging the effectiveness of interventions, so if I can get a margin of exposure from 100 to 500, is that five times better? Is that twice as good? We don't know unless we assume that the relationship is linear. And we don't know if those relationships are linear.

So the continuous approach: This is the one that's been kind of coming forward, I would say, in the risk assessment community. This is where there's been a lot of growth, a lot of good thinking, a lot of research, trying to find ways to come up with continuous approaches; the lowest have continuous estimates of risk.

The chance of liver damage from exposure to this chemical at this level is about 1 in a 1,000, so we could actually come up with those kinds of applications. And, in fact, part of what you're doing there is accounting for population variability, something that all risk assessors have recognized as a shortcoming of our traditional approaches as we don't always include variability both in exposure and in susceptibility as well as we could. And this can now be used

in a lot of those kinds of analyses that we want to do, regulatory impact analysis that the Office of Management and Budget asked the FDA to do.

These kinds of continuous approaches can be used there. They don't fall into that problem that we had with our reference doses where even though we're reducing exposure, nothing good is happening as far as we can -- according to that approach.

Now, limitations here again, data requirements. We need data, and sometimes we need even more and better data. I will say that this is still scientifically controversial. There is still a fair amount of debate in the toxicology, in the risk assessment community about whether noncancer effects truly are continuous or are they discontinuous, meaning there are thresholds and there are levels of exposure that we just don't have to worry about, and pretending that there's risk below those is just crazy.

When we want to do this, we run into some of the problems that Steve identified with cancer risk assessment where we have what's called model uncertainty. We don't know necessarily the right model for doing this continuous modeling, and the model you choose can have an enormous effect on how big you think the risk is. And these are much more challenging to communicate.

Instead of you're above or you're below the safe level, now you've got to say something about what your risk might be, and in fact, you might want to see, say, what is the risk to people on a different part of the variability distribution. These become much more harder to communicate.

So one of the things I want to advocate, looking forward, and this is just some ideas, to me, one of the biggest problems we have here is even if we could decide which of these we want to do is this data problem. What do we have to do to understand the potential noncancer effects of these chemicals? Which tests do we need, what endpoints matter, how can we find points of departure? And frankly, my message is, at the end, we've got to become more comfortable with uncertainty and recognizing that we don't know everything we know, but that doesn't mean we don't know anything. So for noncancer effects, traditionally what is done is to test chemicals for lots and lots of different endpoints. We test them for reproductive toxicity, developmental toxicity, neurotoxicity; we test them in 90-day tests; we test them in two-year tests; there's lots and lots of tests that have to be done. And then from that comes this critical effect that I said, and this is a quote from EPA; it says the critical effect is "the first adverse effect, or its known precursor, that occurs to the most

sensitive species as the dose" changes. And then that's what we use for calculating our point of departure.

Now, one question is, do we have to test things for everything? So this is a study that I did with a graduate student; we took 352 pesticides. Like FDA can do with pharmaceuticals, for pesticides, EPA can require data to be submitted. And there are a lot of requirements for submission of all of these tests that I talked about for pesticides. The data is submitted to EPA; EPA then uses that to do its risk assessment.

Part of doing that risk assessment is calculating one of those reference doses that I talked about. Now, just as -- this is a list of the toxicology tests that are required. The only estimate we could find is that it's about $35 million to do this, and it's probably about 10 years' worth of research for each chemical. The question we were interested in is when we've got all of these tests, we've actually done them, which one do we use to set a reference dose? Which one's important? Are they all equally used, or is it some other thing?

It turns out almost 80% of them use just one of those tests. It's the chronic bioassay; it's probably the most expensive and the longest to do of those, but many of those tests, developmental toxicity, reproductive toxicity, are all

almost never used to set a chronic reference dose. So maybe we can save some time and some little animal lives, doing things a little faster and not doing every test on every chemical.

A second question is we tend to think that when we're looking at noncancer effects, we talk about what endpoint we're avoiding. Oh, this is going to avoid neurotoxicity, this will avoid kidney effects, this will avoid things going wrong in your brain. Well, it turns out, as Steve alluded to in cancer, we've known for a long time that mice and rats do not get the same kind of cancer when they're exposed to the same chemical. And usually, in cases where we know, they don't get the same kind of tumors that humans get when they're exposed. Benzene causes leukemia in humans; it does not cause leukemia in mice or rats. It causes ten other kinds of tumors to increase, but not leukemia. We did a study looking at chemicals from the National Toxicology Program that were tested in mice and rats under identical conditions and looked to see, for the noncancer effects, do mice and rats get the same effects, and the answer is basically no.

The prediction, they get adverse effects in different organs and different parts of their bodies. So the concordance is very, very poor. But what was interesting, the doses are, in this case -- actually the benchmark doses associated with

eliciting those bad things were very similar, meaning that you saw bad things happening at about the same doses in mice and rats, but it wasn't the same bad thing.

The third question is are there other ways that we can predict our point of departure? Because we're talking about how resource and time-intensive testing is, and here I want to go back to some of the things that Steve mentioned. Quantitative structure-activity relationships can be used to make predictions of points of departure. They require past experience with a group of chemicals that are thought to be similar enough to those we want to use. Empirical relationships, I've actually got work going on right now that -- you can predict the long-term benchmark dose, the benchmark dose in a long-term chronic study. Remember the one that I said was used to set 80% of benchmark doses? You can predict that quite accurately from a 90-day tox study; you don't even have to do the 2-year study.

We've got to find ways to more rapidly get tox values into the process so that we're not doing something where we're only focusing on the chemicals that we can find a value for that someone else has done; we've got to have a way to do things faster. And it may be that some of these Tox21 and Risk21 methods will help us, too.

So just to wrap it up, there's a bunch of ways out there to think about doing, thinking about noncancer risk of the components of tobacco and tobacco products. Each has its own strengths and its own weaknesses. And we're always going to have to deal with this situation and find a way to better deal with the situation of not having data -- or actually, what it is, it's not that we don't have data. I should stop using that term.

We don't have authoritative risk values that have been produced by EPA or FDA or EFSA or Health Canada for lots of those values. For most of them, there are data out there. They've been studied, and we've got to find ways to use that data more quickly, more effectively, to inform our risk assessments. And with that -- so I say, we wanted to work on these methods, and this will mean we will have more inherent uncertainty when we're not testing every chemical for every effect, when we're maybe using short-term tests to predict long-term tests, but it will allow us to not act like no authoritative value means no risk.

Thank you very much.

(Applause.)

DR. DRESLER: Okay. Our next speaker is Dr. Annette Santamaria from the Rimkus Consulting Group, and she will be

speaking on Deterministic and Probabilistic Quantitative Risk Assessment for Estimating Cancer and Noncancer Risks of Cigarette Products. I think that's the longest one.

DR. SANTAMARIA: Longest title, not the longest time.

DR. DRESLER: No, no, no.

DR. SANTAMARIA: Well, thank you. Thank you very much. I would like to -- well, first of all, good afternoon. It's a pleasure to be here. I would like to thank the FDA for organizing this very interesting workshop and hosting it. I'd also like to thank my co-presenters for laying the foundation of -- laying the foundation because it's made my job -- it's going to make my job a lot easier, and it may make us get out of here a little bit quicker for lunch.

All right, so it's the right -- it's the left click?

DR. DRESLER: I'm sorry, use the left --

DR. SANTAMARIA: Okay.

DR. DRESLER: Left one to go forward, and then just to move the mouse --

DR. SANTAMARIA: Okay. All right, so this work was -- I'm the Director of Toxicology and Food Safety at Rimkus Consulting Group. We conducted this work with funding provided by RAI Services Company. The opinions in this talk are those of Rimkus.

So we've already heard quite a bit about the application of risk assessments, and what I'll be talking about is actually applying risk assessment to compare estimated health risks between two tobacco products, and I'll just touch upon -- we've already heard about quantitative or deterministic risk assessment, but the focus of my talk today is probabilistic risk assessment. And I'll be walking you through an example in which we used probabilistic risk assessment to compare cancer and noncancer health risks of two cigarette products.

I'm going to skip this; we've already heard it.

So probabilistic risk assessment is used in many fields, including finance, engineering, the nuclear fields, medicine, occupational, and the environmental field. And as opposed to deterministic risk assessment, which uses a single point value such as a mean or a 95th percentile estimate for the exposure parameters and the toxicity values, probabilistic risk assessment estimates the probability of risk based on the full range of model inputs, for example, all the exposure and toxicity values.

And sensitivity analysis with the probabilistic risk assessment allows the risk assessors to take a deeper look into how model parameters actually influence risk probabilities and how variability and uncertainty affects the risk estimates and

which of those particular inputs are most highly correlated with the risk estimates. So what it allows the risk assessor to do is basically identify the most important variables that contribute to risk estimates, so it's a very powerful tool.

It's commonly used to estimate or evaluate estimated health risk associated with exposure to consumer products. There's publications that have done that that are available. Workplace exposures or environmental contaminants, as we've heard about for risk assessment, and it utilizes a computerized software that performs complicated mathematical analyses that can aid risk assessors, and as I mentioned, the identification and characterization and quantification of key parameters that affect risk estimates.

And as we've heard about, the National Academy of Science's 2009 book on *Science and Decisions*, in that book they recommended that a probabilistic framework will provide a substantially more complete quantitative characterization of hazard in that you can quantify, you can evaluate -- uncertainty and variability can be more thoroughly evaluated. It provides for better informed risk management decisions, and risk management options may be more explicit and transparent.

And here are just some recent reports on probabilistic risk assessment or at least uncertainty. On the left is the

EPA 2014 report; it's very, very informative and very helpful if you want to learn further about probabilistic risk assessment. And the WHO International Programme on Chemical Safety also delves into sort of an analysis of the uncertainties in risk assessment and hazard identification.

So what we did was we selected two reference cigarette products to illustrate the methodology, and again, it will just be sort of a broad overview of the methodology. It would take too long to kind of go into all the details, but I'll give you a good illustration of this methodology.

And we estimated our exposure based on the Health Canada Intense HPHCs that we selected for this analysis, and we evaluated and compared estimated cancer and noncancer risks. We actually did it using QRA or deterministic risk assessment and probabilistic risk assessment. And I'll show you the results of both of those in brief detail.

So we used the risk assessment paradigm that I'm sure everyone's familiar with in which our hazard identification consisted of selecting the HPHCs of two reference cigarettes. Our dose-response evaluation consisted of identifying the relevant noncancer reference values and the cancer inhalation unit risk values that were available for those HPHCs, and we conducted, we evaluated chronic inhalation, and we did our risk

characterization to look at those other relationships between the health hazards and the exposures.

For our probabilistic risk assessment, we used Oracle's Crystal Ball to conduct the PRA through Monte Carlo simulation, and we selected ranges for each model input variable, for example, for smoking behavior arranged for the HPHC values and for the toxicity values. And then we calculated cancer and noncancer risks using Monte Carol simulation and run with 5,000 iterations.

And all these variables are up to the risk assessor who is doing this; they can select, they can do 10,000 iterations, etc. But for the purposes of this example, that's what we did and we -- cancer and noncancer risks were rendered as probability plots. So as Robinan showed earlier today, the exposure model, it's based on RAGS Part F, and this is the one that we also used for both our QRA and our PRA.

For the probability risk assessment, as I mentioned, you look at the whole distribution of values and so for -- these are the exposure variables that we used in our assessment, and as you can see here -- is this, is there a mouse? Or is this a pointer?

(Off microphone comment.)

DR. SANTAMARIA: Okay. Actually, I probably won't. Oh,

yeah. Okay.

So what this is showing is that we selected for the likeliest value the mean HPHC yield, and then for a range, we selected the lower and upper confidence limits. And the same holds true for these other exposure variables, such as cigarettes per day. We selected a likeliest value of 20 cigarettes per day, and then we chose a range of 1 to 95 cigarettes per day. The same holds true with exposure duration. We selected 7 to 70 years, and the mid-point between that was 32, so we selected that as our likeliest value. And averaging time is just a function of the exposure frequency times exposure duration. And our daily inhalation rate also went from 15 to 25 m^3/day.

And here are the reference values that we identified for the cancer and noncancer toxicity values, the inhalation unit risks are on the left and you can see the various sources. We used -- and we also started out with using FDA's list of 18 HPHCs that was published in the 2012 *Federal Register*. And these are the reportable HPHCs that are required by cigarette -- or tobacco -- or cigarette manufacturers on an annual basis.

So this is the actual sort of hypothetical HPHC data that we used for our analysis, and on the left is Product A, so you have mean yield, and then we applied the 5% and the 95th

percentile confidence intervals to get our distribution, and Product B is on the right. And so the ones that are highlighted in yellow are those HPHCs that appear to be higher in Product B than Product A.

I failed to mention, but for this risk assessment process, you can actually use this method for comparing not only cigarettes but other types of tobacco products. This is not strictly limited to comparing cigarettes.

So, anyway, the ones in yellow are slightly above the -- or I'm sorry. Yes, they're above the confidence, the 95th percentile confidence interval for Product A, so those appear to possibly be higher in Product B than Product A. The ones in pink are the ones that appear to be lower in Product B than Product A. So with having this data, how do you really know are there any significance -- is there any significance to those differences in HPHCs between these two products if you're comparing them? And so that's when we chose to use quantitative risk assessment to look at those two products and compare cancer and noncancer health risks.

And we've seen this calculation. We used the same one. You actually use the same calculation for both quantitative risk assessment -- or deterministic and probabilistic where you calculate your hazard quotient, which was a function of the

exposure concentration divided by your reference concentration.

And here's the cancer risk calculation where we calculated an incremental lifetime cancer risk based on the exposure concentration that we estimated divided -- I'm sorry, multiplied times your inhalation unit risk.

So for the deterministic risk assessment that we conducted, we looked at the HPHCs that had -- or the individual ones that had reference values, we calculated hazard quotients for each of those, and then we summed or added those hazard quotients to get a composite hazard index, and there was actually -- we saw that there was no difference. It's less than 1% between the two products, suggesting no meaningful differences in noncancer risk for the two products. However, when you look at the individual HPHCs to understand are there differences between the cancer risks for the individual HPHCs, as you can see for -- I've highlighted acrylonitrile and acrolein. Acrylonitrile has a slightly higher -- there's a difference of -- this is -- there's a slightly different -- there's 64% difference between Product A and Product B with respect to the hazard quotient. However, with acrylonitrile, you can see that it only contributes less than 1% to the overall hazard index, bringing into question how significant or how relevant is really that difference between those two

products.

When you look at acrolein, you can see that there are no differences. It's less than 1% in the hazard quotients. However, for both products, acrolein contributes 97% to the overall hazard index. I've highlighted these two because these are the two HPHCs that I will be using to illustrate some of the results from the probabilistic risk assessment.

And the same holds true for the cancer risk assessment. There was a difference between the composite -- I'm sorry, for the cancer risk assessment, there was a difference for the composite ILCR, suggesting that there was a difference between these two products. So this is where we decided, well, let's see how relevant or how important is that difference, or is it really a true difference, and that's where we applied probabilistic risk assessment to look at that.

And as you can see, here's the quantitative risk assessment, the 64% difference between Product A and Product B with respect to the ILCR for just acrylonitrile. And so we -- in summary, for the quantitative risk assessment, there were no differences in composite hazard index between the two products. However, there were -- the composite cancer risk value suggests a possible -- a difference in cancer risk.

So that's when we decided to do the PRA, as I mentioned,

and this is actually from the cover of the EPA 2014 report, and it -- I think it shows very nicely the distributions for the various parameters that go into a risk assessment; for example, concentration in the environment, you can see it has a distribution as does these other variables that go into the model, and we've heard a lot today and yesterday about some of the uncertainties that go into some of these values, and what this distribution does is it sort of accounts or handles some of those uncertainties or accounts for them.

So in our particular example, we chose -- the probability risk assessment model allows you to select a type of distribution for the data that you're inputting into it, and so for the purposes of illustration and for our -- we selected BetaPERT because of the -- what's known about sort of the scientific -- what's known about how people smoke, like most people don't smoke two to three or four packs a day, so it allows you to have most of the data right around the value, what you think is the likeliest value. But again, this is just for illustration purposes, and it's kind of up to the risk assessor which one to choose.

And so here what I'm showing you is our data from our PRA, and again, I'm highlighting acrolein and acrylonitrile for illustration purposes. And what the probability risk

assessment allows you to do is look at -- well, let's look at the mean hazard quotient and compare it to the range of possible values for your comparator product, Product A in this case, and does it fall within that range. And as you can see, for acrolein, the mean hazard quotient falls within the range.

I don't know how to -- oh, there. Okay.

It falls within the range of 5% to 95% for Product A, meaning that there do not appear to be differences between these two products with respect to acrylonitrile noncancer risks. When we look at -- oh, I said, I mean acrolein. Sorry.

When I look at acrylonitrile, when we look at noncancer risks, we have a value of 64.66 for the mean hazard quotient, and you compare this to the range of values for Product A, it falls very close to the 95th percentile, and we'll look at this a little more carefully on a graph to show you a little more detail about that. And the same holds true, actually, for the cancer risks. The mean value for Product B is very close to the 95th percentile.

However, it's important to note that our PRA, just like our QRA -- or as opposed to our QRA, actually, our PRA showed that -- how do I get back?

(Off microphone response.)

DR. SANTAMARIA: Okay. Okay, sorry.

So our PRA shows that the composite hazard quotient mean falls within the range for Product A, indicating there's no difference in noncancer. In addition, the mean ILCR for Product B falls within the range for Product A, again, indicating that there are no differences in composite cancer risks for these two products.

So this is just -- this is the probability plot illustrating our acrolein noncancer health risks, and you can see they overlap pretty much entirely or very much so, and you can see here's the mean values for both of them, indicating there are -- it's confirming there's no differences.

And so for acrylonitrile, in this particular probability distribution plot, we have plotted the 5% and the 95% cancer ILCR values for Product A and for Product B. We have the 50 percentile here in red, and as you can see, and as I mentioned in the table, it's very close to the 95th percentile for Product A, indicating that 50% of the risks overlap, but it's possible that 50% are higher for Product B.

So in order to look at that in a little more detail and to dig deeper into what really is the difference between -- what are the differences between the possible cancer risks between these two products for acrylonitrile, we plotted the upper end of the distribution curve, so here we go for Product A, it's

from -- did I?

Oh, sorry. I didn't -- okay. I think it's because it's on the left, I get confused. All right, okay. Sorry about that.

So here we digged deeper into the upper tail of the distribution curve, and here's the 90 percentile for Product A and 99th percentile for Product A. And Product B is over here, it's 84%, and what that's basically telling us is that approximately 84% of our cancer risk or estimated cancer risk for acrylonitrile falls -- are comparable between these two products, so it's not 50%. And what's happening is that the differences between these two products with respect to acrylonitrile cancer risk is at the very high upper end of the graph where we have very extreme values.

We have in a 90 -- up to 90 cigarettes per day, we have the highest upper -- the confidence limit for the HPHC, saying it's at this end, so it draws into question how significant really are the differences at this upper end, and are they really relevant to assessing and comparing the cancer risks for two products, two tobacco products.

In addition, PRA allows you to do a sensitivity analysis in which you can look at which exposure and hazard model factors lead to the greatest variance and risk, and also to do

a rank correlation, that is, identify those exposure and hazard factors that correlate the most closely with risk.

And for illustration purposes, we selected one example graph, and you can generate these graphs for every single HPHC that you're analyzing. In this case, this the estimated contribution to variance for acrylonitrile, for the acrylonitrile cancer risk. And as you can see here on the left, this is cigarettes per day, and these are our exposure variables and hazard variables that go into the model. And on the left you can see that the contribution of variance is really being driven by cigarettes per day, exposure duration, and assumptions surrounding the inhalation unit risk value that we're using and assumptions about average and time.

Combined, this contributes 96% to the ILCR estimate, whereas Product B -- I'm sorry, where there's the product acrylonitrile yield contributes less than 4%, indicating that most of the cancer risk is really being driven by our exposure parameters and uncertainty surrounding those exposure values and the toxicity value that we used. Did I -- yes. So then when we look at which of these variables correlate the most with our cancer risk, with 1 being the variable is highly correlated with the risk estimate and 0 being no correlation, we can see here that, again, the cancer risk estimate is being

driven primarily by cigarette smoking exposure -- sorry, behaviors. Cigarettes per day exposure duration. And the uncertainty surrounding the inhalation unit risk, and the acrylonitrile yield is not very -- is not highly correlated with the actual cancer risk estimates.

So as I've mentioned, this is purely a hypothetical QRA/PRA in that it was just done to sort of illustrate the methodology and the types of information that you can glean from this type of risk assessment. It was using smoking -- but as with any good risk assessment, we highlight a few of our assumptions and limitations.

Again, we used HPHC yields, and we've all heard that -- as a proxy for exposure, and there's, as we've heard, there's limitations and problems with that, and of course, assuming 100%, so these are all very worst-case extreme estimates of cancer risk and noncancer risk in that these provide sort of a way to sort of compare products, but they are not really generating estimates of absolute risk. If you were to do that, you would really need to use slightly different -- you would have to really understand sort of the parameters that are going into your model and make sure that they have very little uncertainty surrounding them.

So then, as I mentioned, there's variability and

uncertainty that are evaluated in the model regarding the exposure factors, the HPHC, and the yields.

So, in conclusion, what our examples show is that there were no apparent differences in composite hazard index between the products using deterministic or quantitative risk assessment. The HI, the hazard index, was predominantly attributable to acrolein, and our probabilistic risk assessment supported the QRA finding that there were no differences in noncancer health risk between the two products.

With respect to our cancer conclusions, our cancer risk conclusions, the QRA data showed a 34% difference in the composite ILCR between the two products, which was predominantly attributable to acrylonitrile, and however, our probabilistic risk assessment showed that 84% of the cancer risk actually overlapped between the two products, so in reality the differences were really probably closer to 16% at the very, very high upper end of the distribution. And sensitivity analysis suggests that smoking behaviors and uncertainty surrounding the toxicity values represented 96% of model variability, and they were more strongly correlated, also, with cancer and noncancer risks. I didn't show you the noncancer data, but -- then differences in acrylonitrile yields. So the probability risk assessment data suggests that

cancer risk is comparable between the two products.

So my general conclusions for these methods is that comparative risk assessment of two tobacco products may require varied levels of analyses kind of depending on what types of products you're comparing, how comparable they are with respect to their HPHCs, etc., to fully understand the relationship between exposure and health risk.

And deterministic QRA, single point value estimates -- I mean, using single point values or 95th percentile is an excellent way to do sort of a screening and understanding whether there's a possible difference in health risks between the two products. However, probability risk assessment -- probabilistic risk assessment will provide in more detailed analysis the relationships of the exposure and toxicity variables and health risks and really help you understand what's driving those differences before you make conclusions that may not necessarily be accurate regarding differences between the two products. So finally, differences in HPHC yields between the two products may not necessarily equate to meaningful differences in cancer or noncancer health risks.

I'd like to acknowledge two of my colleagues at Rimkus, Marshall Krotenberg and Scott Drouin, who have been very helpful with these analyses. And thank you very much, and it's

been a pleasure.

(Applause.)

DR. DRESLER: Thank you.

We will break for lunch now, and then we will -- after lunch, we'll have one more presenter, and then we'll have the panel discussion. So that line, I understand, can be long at the cafeteria, the restaurant that's up there. Please go ahead and jump in line up there, come back in 1 hour. Okay, so 12:20 we'll start back up again.

(Off microphone comments.)

DR. DRESLER: 1:20, 1:20. Sorry.

(Whereupon, at 12:22 p.m., a lunch recess was taken.)

A F T E R N O O N S E S S I O N

(1:25 p.m.)

DR. DRESLER: Okay, we'll go ahead and get started, and we'll just, we'll have people be coming in from their lunch upstairs. So our next speaker in Session 6 now, which is on Risk Assessment: Mixtures, will be Dr. Moiz Mumtaz from the Agency for Toxic Substances and Disease Registry, and he'll be speaking on Methods for Toxicity Assessment of Chemical Mixtures.

Dr. Mumtaz.

DR. MUMTAZ: Good afternoon. I'm sure you all are ready to go home, and we'll try to -- we all have time, but we try to do it as soon as possible. So I want to thank the organizers for bringing us all together. I think most of the faces of the speakers were those I knew for about last 20 years, 25 years. So it's an experience that we'll remember for some time.

ATSDR, the Agency for Toxic Substances and Disease Registry, deals with unintended exposures and health defects

versus exposure and health effects. So we deal with hazardous waste sites and various chemicals that are found there. I want to say one thing, that if you repeat so many times some things, then we start either getting confidence in them or reinforcing them in others. I think some of the slides you look at, these are just not repetition; these are on purpose there so that we reinforce what we have said before. And Dr. George Gray said that there are data that are not being used, and we had thought about this in 1988, said there are data which are not being used, and so you'll see some of the things which are said today talk about what can we do with the data that are there and what else is possible.

So when I came to U.S. EPA National Center for Environmental Assessment, they said you're going to do mixtures toxicity assessment. I said, oh, that's great; I want to do that because I was dosing animals with various P450 inducers and inhibitors. I said, okay, that's great; that's right what I want to do. Little did I know that it had nothing to do with testing.

So they said you're working on a mixtures project for Superfund sites, and you'll write up a program of research and do something about it in terms of what do we do in risk assessment. So I took President Kennedy's saying and

paraphrased it; ask not what the data can do for you, ask what you can do for the data --

(Laughter.)

DR. MUMTAZ: -- in terms of asking for funding from EPA and the Office of Research and Development. But what George said was that ask not what you can do for the data, because there are no funds available anymore for environmental toxicology, so what can the data do for us? So what can we squeeze out of the data, and these are the two things we came up with. One is derivation of additional dose for target organ toxicity rather than just stay with the -- so what can we do, what else can we do?

And the other thing which a lot of people have touched about are interactions, and what can we do with that. So you have seen this so many different ways. Basically, the whole sense of toxicology is to -- oh.

The sense of toxicology is this: getting the dose-response curve so we can see what data are there and what we can use for. And from there we most often get the lowest observed adverse effect level, and then we try to get no observed adverse effect level. In all of the studies you will stick right there, but more recent studies you will see there's some data down in this area, and then basically use the NOAEL, as

Dr. Gray and others have pointed out this morning, you use uncertainty factors to derive what we affectionately call health guidance values, and each one of us have a different nickname depending on which agency we work.

Okay. As you all know, every chemical causes multiple effects; it's not one effect. And most of the regulatory agencies stick with the critical effect, which Dr. Gray again pointed out this morning is the most sensitive effect in the most sensitive species at the lowest dose possible. And so most, even though we review the literature and toxicological profiles which our agency documents, deliver the whole database.

But ultimately, the focus is on finding that one study or the series of studies that support the study and trying to define the critical effect, and so the reference dose is derived based on the critical effect and these -- the other dose-response curves are forgotten, or they're used as weight of evidence to support what you're doing.

And this is a slide which Glenn Rice showed yesterday morning. There are three different ways we can do mixtures risk assessment. One, if you are lucky, which you are already aware, is you have the mixture of concern tested. There are 25 chemicals which you are dealing with, there have been exposed

inhalation to human beings, and you have the dose-response curve, and you go with that, and you have no uncertainty factors. Those are far and less.

Then there's the similar mixture approach where you have gasoline that can be tested in inhalation studies, but you are dealing with gasoline that has been leaking from an underground storage tank. So it is similar gasoline, the composition has changed, but it's the inhalation study, you look, do some structure activity relationship studies to see the proportions are good, the measure of components to a similar mixture. But more often, as you all have been hearing the last 2 days, we use risk assessment based on the components. And you have seen this so many different ways.

Briefly, you have four components in the constituents of this mixture, A, B, C, D. And you have the reference doses for them, you have the exposure data, you calculate the hazard quotient, you add them up, you convert the hazard index, and if it is greater than 1, then you are concerned about it, and that's how risk assessment has been done.

And in 1986 guidelines, U.S. EPA, this was the basic concept that was present, and of course, this all started from FDA long before -- or at least about when we were born. So the hazard quotient is exposure over allowable level reference

dose, and each agency, as I said, calls it in different ways, and we use that ratio. And if you do that, what is happening is we are using just a critical effect, and in '88, '89, '90 we looked at a lot of data, and we said, well, there's a lot of dose-response curves for -- effects and maybe we should look at, see what we can do in terms of deriving -- oh, that went too fast. Okay. It's working.

So what we said was that each of those dose-response curves can be used to derive specific values for that target organ. We coined them target organ toxicity doses. If we do that, then what we find is if we use the guidelines hazard quotient calculations, all we had was reference dose for three different endpoints, and the rest of the endpoints were not considered because the bulletin did not have any data. They said we don't have any reference value, so we do only for those endpoints.

So in 1989, the Risk Assessment Guidance for Superfund came in, and the EPA said, well, use the item B for all of those, so use the item B across for every health effect, and you can see when you do that, you get hazard indices of different numbers. So we are overestimating the risk here because we're using critical effect. So in 1992 we came with this concept of target organ toxic doses. If you calculate for

each individual health effect, then the characterization of toxicity is much better, it is more realistic, and it will -- of what we want to do in public health.

So this concept, I understand from Vince Cogliano, that EPA is now applying to derive multiple values for multiple health effects so that we can characterize toxicity well. The other point, which is the handicap of the hazard index approach, is interaction that we always did with one chemical at a time, and it is important that we look at what happens again.

We looked at the data at that time, and we found that there are interactions occurring. There are three concepts I want to present here very quickly. One is -- antagonism is where if you have two chemicals which are causing some toxicity when you add them up, they are less than additive, and the other one is where you add two chemicals, and it is more than the additive, and this is something which would be concern, and this is something we take advantage of pesticide toxicology where we mix two chemicals and try to knock out the insect as well as kill it, and that's the synergism which we use in toxicology.

And so there are also chemical interactions where one chemical is totally inert, but when you put in the system, it

causes more than additive toxicity, so these are the three basic concepts we try to look at the data, and there's a lot of information, actually, in terms of interactions and where they can occur and how they can occur.

So as a chemical goes through this process of being exposed to and crossing a membrane, whether it is through oral exposure, through dermal, or through inhalation, it has to cross the membrane and get into the blood or lymph where it is circulated all across various organs, and some of it is metabolized. Either it could be activated or deactivated, and some of it is stored in various organs, such as bones and fat, so forth, where metals and so forth get into the bones and PCBs and all fat-loving chemicals get into the fat. And then the chemicals are distributed, and then they are ultimately excreted through various sources. So this is for one chemical. If there are multiple chemicals coming through, there's a possibility of interaction at every level, and there's a lot of data in the literature which supports that there are interactions occurring at various levels.

So when we first looked at it, we thought we can look at the hazard index approach in a little different way that we can look at calculation as we are doing for each individual chemicals, and then we can look at interaction, see what we can

do either qualitatively or quantitatively. And that was the idea of coming up with a hazard index which is adjusted factoring in the role of interactions.

So we set forth in late '80s to look at how we can use this interaction data, and we looked at the best way to judge and how we can empirically come up with a method and how that can be supported in terms of plausibility of toxicity.

And when we were listening today to talk from IARC, we actually use the IARC method to use the weight of evidence and see how we can use the interaction data as -- and support by weight of evidence and how can that methodology be consistently applied, and is there any way we can quantitatively adjust a risk assessment, and finally, how we can validate what we are doing?

And so the weight of evidence first deals with qualitative assessment and then set it up in a matrix so we can look at what is going on and adjust quantitatively, and very rarely we have done that even though the methodology which we published in '97 and -- at EPA, there are ways to adjust quantitatively, but the data clears so much variability and so much uncertainty that we have more often used qualitatively how to deal with that.

And so if you look at -- this was the ideal situation

where we could categorize the data completely first. The bottom line, whether the interaction is additive, less than additive, or more than additive, look at the mechanism or the mode of action to see if it is for the chemicals themselves or for structurally related chemicals. So if it were Chemical A and B which you are interested in, it would have been one, but this is for related or structure-based chemicals, and then we know the toxicological significance of it in terms of what it is doing, and then we had other parameters which play a very important role in terms of route and what kind of data they are. But as we started applying that more often, we realize we don't have enough data for the latter part, but the first three we can definitely provide some information on the interaction, whether it's less or more. We can look up the strength of the mechanistic data and also look at how -- what kind of impact that interaction has.

So this is the summary of the weight evidence scheme which we published way back in '92. It looks at whether there are data which are available which we can use, like -- and then if there are data which are available from related compounds or there are inadequate data. And this pretty much parallels what we saw, strength, which was strong evidence versus weak evidence versus inadequate data. So we planned the same

strategy of sufficient confidence and then limited versus inadequate.

And so each weight of evidence, and there are several of them which have been published now, you can take a summary, a one-page summary of all the interactions that are available for the two chemicals. It gives you the direction of interaction, it gives you the mechanistic information, it tells whether it has an impact or not, and various modifications and limitations, and always risk assessment deals with limitations and uncertainties and provide the references.

So if you look at this scheme for metals, which actually have the best data, you can see that effect of lead on zinc. It's additive; it is based on certain criteria which is not exactly the chemical formulation or the sort we are looking for and what kind of information we have.

If we look at effect of zinc on lead, for example, we know it is less than additive. We know the mechanism, and we know that is -- it impacts the toxicity, and we know the endpoint, so this could be for various endpoints we can determine the toxicity, and the detail of this explanation can be a paragraph in terms of what we know about the interaction of zinc on lead to cause less than additive toxicity, as shown here.

So once we do this kind of analysis, we can set up metrics

like we saw yesterday for weight of evidence and look at, see whether there are multiple binary interactions occurring that would cause more than additive toxicity, then you have to be a little bit more careful. If there are multiple interactions occurring but the toxicity is less than additive, then your hazard index approach could be a predictor of public health.

So to look at that, additives, they all have all of these interaction profiles. The purpose of the interaction profiles is to look at database of chemicals that are found at waste sites and evaluate the toxicity of individual chemicals and the interactions data and then make a recommendation to the public in terms of what has been done. So in a way, being at the Department of Health and Human Services is a little bit advantage that you don't have to go to the -- and quantitatively determine how far the levels have to be done. We are more advice -- when we go to the community, we explain that you are being exposed to these chemicals, but the toxicity is not what you're looking at.

So these interaction profiles, a bunch of them are on the website, and you can see several chemicals we find in cigarette smoke are there, and we have to go into each of those profiles and look at what kind of interactions are documented, and you'll find there's a lot of information that you can use for

qualitative purposes. And that alphanumeric which you just saw can be simplified for public, as shown here. So you don't have to show the alpha Ab-2 and all that to the public. You can just be very simple and explain that some of these interactions. This actually is a mixture found in corn where people are using pesticide and herbicide and fertilizers, and you can see that some of these interactions have high confidence, some of them have medium confidence, and some of them are not so confident, but you can see how each of these chemicals have been documented in terms of interacting with another peer.

And so for public, you can explain that we have high confidence that this is what is happening, that the effects are additive, they are not synergistic, and the levels which you are being exposed to are below these levels. And same thing for researchers, this helps point out where the data gaps are and people can perform research.

So getting to actually we have so many models, so many methods available, but when it comes to site-specific assessments, particularly as we are exposed in emergency situations, there's no time to do further studies and more data analysis. We have to implement what we have. And most often, this is a challenge we face from time to time. And the

mixtures guidance which we have at ATSDR, at EPA, at NIOSH, Health Canada, they all have the basic same principles; they allow you to do a risk assessment, but they give flexibility because they are data driven. And in future, and I have not seen really data which we can use, currently use, but weight of evidence can be supported by high-throughput data which are being now generated through ToxCast and Tox21.

However, there's a lot of shortcomings in that, in terms of the X and Y being used, the ratio that's being used, and the target organ doses which are being used in those assays are very critical. And so it is essential, then, we have to look at what the data are saying in terms of in vitro, but it is a cleaner system hopefully, less noise, and as people after me speak about uncertainty and variability, maybe it will be less; we'll have to just look at the toxicity of what it means.

And, of course, from public health perspective or for you at FDA to develop new products, it is important we communicate what the uncertainties are and how we can develop better product by eliminating certain chemicals. And I think one of the important things in risk assessment in general is the harmonization of methods. There are so many methods being developed, and we have to, in terms of measuring the chemicals, in terms of doing risk assessment, that we have to harmonize

the methods so that cigarette smoke, I see so much data coming out of China on smoking. So we'll have to have the same methodologies available in terms of analytical tools, as well as risk assessment tools that can be used. And I think it's very important that we should not be looking only at cancer risk even though -- is everything okay?

In terms of smoking, we always look at cancer, we're looking only at death, but the fact of the matter is we have other health effects that are caused by cancer which has a latency period of, you know, a long latency period, 20, 25, 30 years, whereas all kinds of upper respiratory problems start occurring in a few months to years and so is also true of developmental effects, reproductive effects, so we should characterize the toxicity of the chemical and the mixtures so that we get a better view of what is going on rather than go for the cancer and death.

And I thank you for your attention, and we'll wait for questions as they come. Thank you so much.

(Applause.)

DR. DRESLER: We'll have our panel discussion, so could I please have Drs. Roberts, Gray, and Santamaria, and Lynne Haber.

(Pause.)

DR. DRESLER: So raise your hand if you want a white card to put a question down and --

(Pause.)

DR. DRESLER: Okay. So let's start. Let's see. How can we integrate what we learned yesterday in Dr. Asgharian's talk about inhalation dosimetry into the application of existing environmental risk values such as IUR, RfC, etc.? So how do we integrate what we learned yesterday from Dr. Asgharian's talk about inhalation dosimetry?

DR. HABER: I'll take a stab at that. So one thing that really struck me from Dr. Asgharian's talk was the implications of the particles coming together in the high concentrations where they exist in the cigarette smoke, and his comment that MPPD, which is kind of the go-to modeling approach used for environmental risk assessment, doesn't predict well the particle deposition in the respiratory tract for under the cigarette smoke conditions, and so I think that's something that needs to be considered. And the extrapolation from the -- when we're using environmental risk values such as these inhalation unit risks or the RfC and applying it to the use of tobacco products, there are approaches that have been used in looking at differences in dosimetry, so just as we do dosimetric adjustments from the animals to humans, you can do

similar ratios of looking at the differences in deposition in humans under the conditions that were assumed from under environmental risk assessment versus the deposition that would occur in humans under the smoking conditions or whatever tobacco product exposure conditions are used. So I would recommend that that be done for any sort of refined assessment for tobacco products.

DR. DRESLER: Anyone else?

DR. MUMTAZ: I think it's also important to realize that in gaseous phase, there could be reactions taking place, and new chemicals might be forming which we are not aware. So there could be chemical, chemical interactions, and that part should be considered when we are doing the risk assessment, not just knowing what we know about, but the fact that there could be chemical reactions taking place in the lung.

DR. DRESLER: And this is for everyone. What data and methods are available to evaluate if the hazards identified represent the total risk from exposure to a tobacco product? What data and methods are available to evaluate if the hazards identified represent the total risk from exposure to a tobacco product?

DR. GRAY: That's actually a really, really hard thing to do. And there have been a few attempts to do this, so the idea

is could we take the components that we know about in smoke, assess them as we have talked about today, and then really what you could do is then compare it to the epidemiological data that says is this what we saw happening in people.

Two problems with that: One is that the attempts that I'm aware of that have tried to do the reference dose approach, part of it, it doesn't give you an estimate of risk that you can compare to the epidemiologic data.

On the cancer side, we know that, as Steve said, our tools for estimating cancer risk are designed to overestimate rather than underestimate risk so that we know that they're not necessarily -- they're used -- they grew out of using risk assessment to be protective, not using it to be predictive, and because of that, they are not ideal for trying to compare to epidemiologic data either. So ideally, what we would love to do, right, you want to check your work, we would love to have a way to say I want to assess the risk of all these things and then go find an independent measurement that said did those things happen, and that's what epidemiology could do interacting with risk assessment. Right now, for a variety of reasons, it's a very hard thing to do.

DR. MUMTAZ: I think risk assessment has to be transparent, and unless consensus is developed, we can be

performing hundreds of risk assessment for the same data, and each one might come with a different conclusion, so it's important that we agree on how we are doing things.

And in general, most of the agencies, including ATSDR, when we do a risk assessment, we send it out for public comments and leave it for 6 months out there for anyone to comment on, and then we look at that, the comments, and make sure we have addressed them. Then we get to a panel of independent scientists to review it, and that's how we try to do some consensus building before we make a final decision on it for public health and for regulatory toxicology. It's pretty much the same process.

DR. SANTAMARIA: I just have a quick comment. I think it really depends on what your goal is. If you are trying to, with respect to how to select which HPHCs, I think for the purposes of comparing tobacco products, you need to start with a subset, and so if you use the HPHCs, the 18 HPHCs that FDA has identified, it will -- those are all you really need to compare those two products for risk, for estimative risks.

I think it would be very cumbersome and very difficult to input all 5,000 chemicals that might be present in cigarettes into a risk assessment, but I think for the purposes, if you're doing a comparative risk assessment, I think it's very

appropriate to start by comparing the 18 HPHCs that FDA has suggested should be used.

DR. DRESLER: Dr. Roberts.

DR. SHIELDS: Yeah, I was just going to point out that we do have an estimate of what the risk is from all the constituents, and it comes from the epidemiological data. The problem is that we don't have any granularity in those data to say, to be able to pick out different constituents or different kinds of products and see how much they contribute to the total risk. So, in one sense, we have sort of the risks that come from the sum, but what we don't -- what we're struggling with is having tools to predict how that risk might change if we manipulate constituents or new products that have different constituents.

DR. HABER: I just wanted to add to what Dr. Gray said, that we don't have the ability to do quantitative analyses for noncancer endpoints, but we can do qualitative analyses, and so a key question I have, and I have not studied the list of HPHC chemicals recently to look at this, but things like cardiac toxicity, like developmental toxicity, is that what we would expect based on that list of chemicals, and if not, what are we missing? And does this relate to questions, like what a couple of people have noted, as far as looking at peak exposures and

not just time-weighted average? Could there be something that's going on that developmental toxicity is not the critical effect?

But when we've got a short-term very high concentration that's having an impact more than what we would expect is something that -- is there something we can get from this approach that Dr. Mumtaz was just mentioning of the target toxicity dose of looking at separate effect levels for different targets? Can that also help us see what's going on for different endpoints? So those are some additional tools that could be used.

DR. DRESLER: Okay.

This is for Dr. Santamaria. So you had said that the cigarettes per day are the strongest predictor of risk, and so if there's a scenario in a smoker who switches from Cigarette A to Cigarette B and that Cigarette B is a no-nicotine product, and they decrease their consumption of cigarettes per day by 50%, what is the risk?

DR. SANTAMARIA: Well, without actually putting that data into my model, I can't really definitively answer that, but if we're comparing an individual that smokes less cigarettes to -- is that --

DR. DRESLER: Um-hum. So because the Cigarette B is the

no-nicotine product, so they decreased their cigarettes per day by 50%.

DR. SANTAMARIA: Right. We would still have to put that into the model and to see where, then, would the mean values fall and where would -- would it be within the 5th and 95th percentile for the original Product A, and then we would look at what the contribution is to the overall risk. Sorry, I'm not sure if I'm answering that very well, but it's kind of hard -- it's very hypothetical, and I'd really have to sort of look at --

DR. DRESLER: I think what they're saying is the cigarette risk went down or the cigarettes per day went down 50%. I think that's what the question was. So does that decrease the risk?

DR. SANTAMARIA: For that particular product?

DR. DRESLER: So between A and B, so if you switch from A to B.

DR. SANTAMARIA: Oh, I see. Okay, what it will do is change the distribution of risks. I'm not sure if it will necessarily change the mean risk for that cigarette. Sorry, I need to -- I don't know the answer to that. I'd need to look at the model.

DR. MUMTAZ: So you're assuming that those two brands

have -- they're only looking at certain chemicals, you know, which are nicotine only, but each tobacco might have other chemicals. There are multiple chemicals in tobacco, and we're assuming that when this person is switching to this brand and reducing 50% nicotine, that doesn't mean he is reducing all the chemicals which are present in tobacco leaf, so there could be cadmium, there could be things which could be cumulative, and it could cause kidney damage, versus nicotine and health effects caused by that. So we have to look at the whole composition of what this person is being exposed to before making the final judgment that --

DR. SANTAMARIA: Just by -- yeah, by splitting it in half, if they're smoking 50% less, I'd really have to just kind of look at how the model would handle that.

DR. DRESLER: Okay.

DR. SANTAMARIA: Sorry.

DR. HABER: Is it okay if I ask a question of the fellow panelists?

DR. DRESLER: Sure, absolutely. Please.

DR. HABER: So I think I must be missing something. If, based on your equation that the -- and so I'm going deterministic, start it with the simple approach, and then we can get fancy. If exposure dose is linearly related to the

cigarettes per day, then you cut cigarettes per day in half, and we've got risk is the linear function we're assuming for the cancer? So if you cut cigarettes -- if you cut cigarettes per day by 50%, wouldn't we expect risks to go down by 50%?

DR. SANTAMARIA: Yes, you're right. I'm sorry. It was clearer than the way that question was worded, absolutely. Thank you for clarifying that.

DR. DRESLER: Okay, thank you.

What types of mechanistic models would be appropriate for hazard characterization in tobacco risk assessment? So this is one we had earlier that we said we'd hold back to this one. What types of mechanistic models would be appropriate for hazard characterization in tobacco risk assessment?

DR. SHIELDS: Yeah, I think there's a lot of interest in developing mechanistic models and mechanistic-based approaches to risk assessment and understanding, you know, how chemicals produce the effect, understanding the sequence of biological events in places that lead to an effect, whether it's cancer or some other effect where the critical points in those steps are -- that the chemical can intervene, modifying factors and so forth.

And then if we understand the mechanism, then it really gives us a very strong basis to do the risk assessment, and it

also gives us some tools that we don't have right now. Right now we're basically giving the chemical, looking at the response, and then maybe trying to figure out a mechanism that explains the response. If we understand mechanisms by which adverse effects occur, adverse outcome pathways, those kinds of things, it gives us the ability to develop models that we can then, you know, for example, do in vitro tests, those kinds of things, and see at points at which chemicals can influence those key events and therefore produce risk of that effect. The difficulty is understanding all of those mechanisms; that's the tall order.

I mean, if we -- once we get to that point, it really does change the way that we do risk assessment, and we can think about if we knew the whatever it is, 100, 500, 1,000 most common mechanisms by which chemicals produce disease, then it becomes a relatively straightforward manner to begin to screen chemicals which -- that can produce effects by those mechanisms. It's just getting to that point where we can have those mechanistic models is the tough part.

You know, we spend a lot of time trying to understand mechanisms for carcinogenicities, chemicals with some progress, but I think there's probably -- and there's so much we really feel like we've got a pretty good understanding of what those

biological events are, what those pathways are. But there's a lot of work to do before we can feel like we've got a comprehensive picture, a library of adverse outcome pathways that we can then use to address risks.

DR. GRAY: I think this is actually a really important thing to be thinking about in this context because all of the uncertainties that Steve just laid out about how we try to understand what's happening in terms of a dose response or a hazard characterization for the components of cigarette smoke. I mean, in the past, the approach has been sort of the linear approach that we saw up here that has been used for a long time.

We know that for certain modes of action, that is probably appropriate, and for some others, it's not. The reason that's important is that in the kinds of questions that FDA is going to be facing here that involve comparing risks, we need to know, we actually need to have good models that predict what is likely to happen when the constituents of a product or the smoke or some other emissions from it change.

Right now we can only tell you what are kind of protective estimates of what could happen that we know are more appropriate for some chemicals than for others. So I think that there's a real need to marry this knowledge that's coming

about how chemicals interact with the body and cause adverse effects with ways that we can predict those effects in a serious way. One hopes that there are tools that are coming that will help us do it. It may well be that there are only a few big categories of different kinds of dose-response relationships, so it won't be as hard as it could possibly be.

But getting to a point where we can be predictive in our risk assessments is really important for making comparisons that we believe are actually providing a benefit to public health.

DR. MUMTAZ: So when we go to our physician for a particular ailment, they still ask some general questions. They look at your body temperature, blood pressure, and ask some questions which are really simple. They're looking at the whole system, not just that one effect.

So the systems approach is something risk assessment is gravitating towards, and the Tox21 testing and all that would give us some idea in terms of the weight of evidence that can be used through in vitro studies to define what that adverse outcome pathway is, and hopefully, that would be the ultimate, maybe not in our lifetime but later on, the future toxicologists and public health systems will be able to do.

DR. HABER: And just building on that, I think that's one

where one of the aspects of the adverse outcome pathways of them being modular becomes really important because you can look at how different chemicals feed into the same pathway and how different pathways interrelate, and so that can help us become more predictive as things develop.

DR. DRESLER: Dr. Roberts, this goes back, perhaps, to follow on what you were speaking just a minute ago. In what ways may the route of exposure influence the development of a TTC? So the route of exposure, how will that influence the development?

DR. SHIELDS: Well, again, back -- as I mentioned earlier, basic precept in toxicology is that the toxic dose can be different by -- depending upon the route of exposure, and so if you -- and this was specific about TTC, so yeah.

So I think if you're trying to develop a threshold for toxicity and you're considering, you know, a body of data and safe doses to try and decide where to set that threshold, you have to be mindful of what -- of the route of exposure and looking at what the safe doses are for the route of exposure that you're intending to address.

DR. DRESLER: Anyone else want to -- we're okay with that? Okay. I have a question here on uncertainty and variability, but since I know that's -- yes?

DR. HABER: Yeah, I just wanted to note that there are publications that have developed TTCs for the inhalation route. I'm a little vague on them, but my recollection is that they did not consider dosimetry issues, so there might be some improvements that can continue to be done, but there are -- and I don't think they're quite as well and broadly accepted as the oral TTC values, but I know that this is an active area of research, so people don't have to start from scratch in developing TTCs for inhalation.

DR. DRESLER: Okay. The question that I have for uncertainty and variability, I'm going to save that for the next session, okay, because it's on that, so I won't forget that question, whoever asked that.

Several speakers have noted the importance of considering peak concentration, not just TWA. How does that impact the methods we have seen today? So the importance of considering peak concentration, not just TWA, how does that impact the methods we have seen today?

DR. HABER: I think that's really a critical issue that I've been thinking about a lot and just watching the talks, and one thing that really struck me with Dr. Santamaria's talk was in the acrylonitriles, for example, that that was the driver.

In my talk yesterday where I showed the importance of the

dose metric and showed that for -- the example was actually an acrylonitrile trial and the differences for looking at the dose metric. The metabolite there, which is thought to be causing the tumors, and there the example was that peak concentration is actually much more predictive of the dose response than time weighted averages.

Well, if the inhalation unit risks that we're using is based on time-weighted average, that might be health protective for the environmental exposure, but we need to go back and relook at things, then, for something like cigarette smoking where you are getting these peak concentrations, and you may be vastly exceeding the concentration that's causing tumors in the human population. And one thing I was realizing is we don't tend to have that sort of data, so addressing this sort of thing may -- you know, the in vitro testing may help, but we may need to be using different sorts of testing paradigms to really address that issue and get information on peak concentrations to the degree that it's needed for these specific problem formulations.

DR. SHIELDS: Yeah, just a general point. I think the dosimetry and what the proper dosimetry is for a response is really important, and sometimes maybe we don't pay enough of attention to it, but sometimes we use dosimetry of convenience

rather than maybe what's correct.

I mean, it could be the peak concentration, could be the average concentration, it could be average concentration times time; time is a factor in there as well.

So there's lot of different ways that you can express dose, and sometimes, you know, we have need to think about what is the most meaningful way to express dose based on the chemical and what it's doing in its mode of action and those sorts of things. And then just on a technical detail sort of an issue is the toxicity value. I mean, some assumption about the proper dosimetry was made in coming up with the toxicity value, whether it's a reference dose or a cancer slope effect. And so we have to be, you know, sort of mindful to -- when we're coming up with an exposure estimate, that we don't create a mismatch between what we're measuring and how we're expressing the exposure versus how exposure is expressed inherently in the toxicity value.

DR. SANTAMARIA: I just have one quick comment. And again, I think it kind of depends on what the purpose of your risk assessment is. If you're trying to evaluate absolute risks of a product, then yes, I think you'd have to consider these issues if you're doing a comparison of two products.

If you assume they're both continually exposed at peak

levels and you compare them based on that, then it won't really matter if you're comparing peak levels or TWAs for two products. You really need to do fit-for-purpose risk assessments, and that will impact how much detail you get into, mechanisms and things like what Lynne brought up.

DR. DRESLER: Dr. Santamaria, this is for you. Why did you choose to run only 5,000 Monte Carlo iterations rather than the more standard 10,000?

DR. SANTAMARIA: Just for purposes of illustration, this was sort of our main reason. And we would've gotten the same results. No, they could have been slightly different if we had run 10,000, but this was purely to illustrate the methodology, so it was sort of just we chose that for use.

DR. DRESLER: Okay. Dr. Gray, so this is -- this is a lot of words on this card.

(Laughter.)

DR. DRESLER: If we assume no risk below an RfD/RfC, do you think we could develop and use a slope factor to more accurately estimate increases in risk above the RfD or RfC? This may be more easily done for PODs derived via benchmark dose analyses.

Make sense?

DR. GRAY: Yeah, sure. That's a great idea. I mean, so

the question is, is if -- even if we assume that there is a population threshold, then our reference dose or something like that is a reasonable estimate of that. Could we come up with a way to estimate the risk above that? And I think the person asking the question is exactly right; something like a benchmark dose or some other kind of dose-response modeling would be very helpful there. Except we never, almost never have exposures that are that high and -- oh, in fact, I was wrong. The reference concentration has still got that safety factor involved for a reference dose so that, in fact, what we would want to be is about a hundredfold above the reference dose before we're even in the range of our dose-response model. So this is still going to be hard, so you remember that our point of departure is then divided by a factor of a hundred or more. Dr. Mumtaz had a really nice graph of what's happening there.

So there's a big space between the reference concentration and our observed data that still makes it really hard to say what might be happening in that region. Once you get up near benchmark dose, we can estimate that kind of risk, but exposures don't get that high very often in most cases.

DR. DRESLER: Okay. One last question, so I think --

DR. GRAY: Lynne thinks --

DR. DRESLER: Oh.

DR. GRAY: Lynne thinks -- Lynne's going to straighten me out.

DR. HABER: So I'm not going to straighten you out. I think you had a lot of good comments. Certainly, I also wanted to note that the method that Weihsueh Chiu showed yesterday is useful for addressing that which basically can account for -- it's usually uncertainty factors, so it's not just assuming continuous dose response all the way down to zero.

And also a little bit of a plug, the website that I mentioned yesterday on the dose-response framework, chemicalriskassessment.org, has a number of case studies that were specifically designed to address this issue of how you estimate risk above the RfD or RFC in response to the science and decisions. It clearly is a very important problem for risk-benefit analysis, and it included a case study example of the type that was mentioned, as well as some other approaches that use more mechanistic approaches.

DR. DRESLER: Okay. Also, Dr. Haber, you had said that during the break that you wanted to answer one of the questions from yesterday that you had thought of. Did you still want to do that while we have -- no? No, if you're okay, I'll ask my boss.

DR. HABER: So yesterday I was asked about the apparent mismatch between the issue of inhalation dosimetry for gases being based on assuming a steady state and what are the implications for intermittent exposure. So the short answer is that if you have not reached steady state, then from the perspective of the human exposure, we're being health protective because the exposure is lower. Where I think there is still a data gap and one of the challenges is that in the acute animal testing, we may not have reached steady state, and that's not really accounted for in the development of acute exposure limits, so I think that's something that additional research is needed, Dr. Woodall.

DR. DRESLER: Okay. One last question which may be short. Dr. Roberts, you mentioned or implied in your talk that some carcinogens have thresholds. Please explain or justify.

DR. SHIELDS: Yeah, I think there is a concept that some carcinogens have thresholds and that there are some critical events associated with the carcinogenicity, and if those don't occur, cancer will not occur, and that if we can quantify the -- identify where that threshold is, we can develop safe exposure levels, that sort of thing. And EPA considers threshold carcinogens -- there are some carcinogens and carcinogenic effects that are managed or assessed on a

threshold basis. The default position is a linear non-threshold approach. But if sufficient data are available to justify based on the presumed mechanism of carcinogenicity that a threshold exists and that can be identified where it is, then it can be managed that way. So we don't always assume that every carcinogen is a linear non-threshold.

DR. DRESLER: Okay, all right. Thank you.

So I'm getting used to that red light which I usually don't like to do, but anyway, I would like to thank the speakers and the panels for another great discussion. I had a few more questions, but -- so I'm sorry if I didn't get to them yet. Thank you very much for your presentations.

(Applause.)

DR. DRESLER: We'll take a 10-minute break, okay? So maybe coffee, but let's start back in 10 minutes, okay?

(Off the record at 2:21 p.m.)

(On the record at 2:33 p.m.)

DR. DRESLER: We'll go ahead and get started again, please. Okay, good. Okay, so there -- thank you. There was one question that was left over that the questioner really wanted to ask, and so Dr. Santamaria was kind enough to say she would come up and answer this one specially.

So the question was how does your PRA distinguish between

input variability distributions and model or data uncertainty to calculate the risk and confidence intervals on that risk?

So, Dr. Santamaria, if you can --

DR. SANTAMARIA: Luckily, that one is pretty straightforward and easy for me to answer. Our particular PRA did not account for differences in uncertainties and distribution. You can do that. You just need to know what the magnitude is for those uncertainties and for the distributions. And you can use PRA to do that and get a more refined estimate of the distribution of uncertainties and the magnitude of the uncertainties and how they contribute to the estimates of risk.

DR. DRESLER: Thank you.

Okay, so now to the last session, Session 7, which is on Risk Assessment: Uncertainty and Variability.

Our first speaker will be Dr. Richard Hertzberg from Mathematics Consulting, Emory University, speaking on Uncertainty in the Risk Assessment Process.

DR. HERTZBERG: Okay, thank you to FDA for inviting me to talk on this interesting issue. My background is in biomathematics and biophysics. I've been -- 25 years at EPA working mostly on mixtures risk assessment guidance and approaches, but I decided, since this is the final session, that it would be easier to take the 30,000-foot approach. I

mean, you've all been very good and dedicated for a day and a half now, focusing on some very difficult, pithy, intense issues, so I'll give you something that's a little more relaxed.

What's really advantageous about being in the final session -- and kudos to all of you who are still here -- and talking about uncertainty is that Ken Portier and I are in this incredible position of power, that after hearing all of these wonderful presentations the last day and a half, we get to pick them apart.

(Laughter.)

DR. HERTZBERG: But seriously, I mean, the election is over now, so I think it's time to return to civil discourse. So we're going to try and be a little more evenhanded and not so -- not pointing any fingers at anybody. Let's see. It didn't do it. I'm on the right one.

DR. DRESLER: I'm sorry, the left one is to go back.

DR. HERTZBERG: Oh.

DR. DRESLER: I'm sorry. The left one.

DR. HERTZBERG: Oh, okay. Go back.

Okay, here's the overview. I'm going to start with just some easy examples from extrapolation. Many of these in the concepts and examples have been mentioned or alluded to in the

past day and a half, so it's going to be easy to, I think, stick with my time frame that I'm allotted. Then I'll talk about three different categorizations of uncertainty that I came up with. They're not original, not terribly original.

The first one you heard Dale Hattis talk about, but not quite in the same words. He likes mine better because it has more syllables. You'll understand when you see it. And then talk a little bit about research areas in the future for addressing and possibly reducing uncertainty in the risk assessment process.

So you're seen this before. I'm just going to point out one big difference why the 1983 NAS paradigm is not a good approach, especially for mixtures. It's because we always are resource poor. Too little time, too little money, too little people. Little people? Too few people --

(Laughter.)

DR. HERTZBERG: -- to get everything that we would like to do and evaluate it the way we would like. And so it's really important to match your dose-response assessment with the exposure assessment, do them hand in hand. Then you're not going to be gathering inhalation data when you really want an oral exposure or high dose versus low dose or intermittent versus continuous. So you want to try and match that and

optimize what's going on.

I had to use this as my example for extrapolation. If you can't quite read it, it's a little fuzzy. The dog is pointing at the three pirates, saying one, two, three, and then he counts himself as four meat eaters. He gets to the grill and counts one, two, three hamburgers. At the end he says I always check their math; their math is always wrong. Just like Eeyore, same tone of voice. This is not really to poke holes in quantitative approaches. I mean, that's my bread and butter and what I was trained to do. But it's really that extrapolation is all about what you have versus what you want.

So here are some examples, and I'm just going to talk about the last one. You heard the first one talked about several times. I think Dr. Haber did a great job of kind of transferring a repeated exposure into a continuous exposure. Dr. Santamaria expanded on this, too, and showed how you can even do better than that with pharmacokinetics and other concerns. We've heard the others talked about in various presentations, as well.

The final one is interesting because if you think back to what Dr. Mumtaz was saying, there are ways to not only characterize what we know about the binary interactions, but also how relevant we think they are to human toxicity, the

quality of it and ways to quantify it. In the interest of time, he did not show you the quantified approaches. It turns out he had time, but we didn't think so when we were working this thing through the first few times.

EPA has a way to modify the hazard index to incorporate the ATSDR approach. EPA has an approach, too, which adds more uncertainties. Which approach do you take? But when you end up with this mixture assessment even as a modified hazard index, you have to have this combination of judgment of a structured weight of evidence framework and some modeling for either the exposure or the dose response. And it's this combination that gives you some insight into what the mixture might actually be doing in terms of posing a risk. So it's not a simple answer. In fact, there was one quote, I think it was H.L. Mencken who said that all complex problems have a simple answer, and it's always wrong. So I'll just leave you with that.

This is one of the examples I liked to use in the class I taught at Emory, is to show you that you have various ways you can get from what you have to what you want, and this is an easy example of the rat inhalation study to a human oral estimate of risk. You can either go rat-to-rat inhalation to oral, or rat-to-human inhalation-inhalation, and then do the

final step. And you really should make this decision based upon what you understand about the mechanisms and the similarities across the routes or across the species.

We know so much about the differences in architecture of the respiratory tract in rodents versus -- especially rats versus humans. It's probably not a good idea to go rat inhalation to human inhalation. It would be much safer to go across this way. And then, of course, the ideal would be just to have it all modeled mechanistically. We heard about this question in the last panel discussion. If we had enough mechanistic information, we could just go directly from rat inhalation to human oral and handle all of those other vagaries and uncertainties about peaks and averages.

This is an example because it's very much in the public eye. One of the concerns about clean power in the Obama campaign has been to reduce the use of coal-fired power plants, and when you do this, you have a concomitant reduction not only just in carbon but in particulate matter.

And so when you have in the case where I live, in Atlanta, very high heat, high humidity for a long portion of the year, which can be anything more than 2 weeks, actually, it's about half of the year, and you have asthmatic children who can't go outside and play because the combination of heat and humidity

and particulate matter is so strong that when they're already asthmatically challenged, it's just too risky to go out and try and do some exercise.

How you would scale this, how you would model this to go from a healthy person to an asthmatic child is very, very difficult, and if someone knows how to do this, I would love to incorporate that in my next presentation somewhere, of how to handle mechanistically this very difficult challenge.

So the other thought I had was to come up with some ways to categorize uncertainty, and this might be of interest not only to explain to you how different people address the concept of uncertainty or how they interpret it, but perhaps give you a way to think about where uncertainty arises and how you might want to take action to reduce some of it.

And the first idea here, this is called aleatory uncertainty, it's the natural variation that cannot be reduced. And we can do the arguing of the -- sure there's some uncertainty equation in -- and all the other examples of are we always perturbing the human system, and so even in epi studies, we can never really know what's going on.

Or is it just that there is natural variation in biological systems, and we cannot know enough to handle it all as you get into personalized medicine and more on population

characteristics and get into what EPA calls cumulative risk to make it very, very focused and population-centric? We'd like to add more and more information on this, but there's a certain point we just can't add enough to handle all the variation. So more knowledge gives you a better sense of population variance, but it doesn't really reduce it down, where lack of knowledge is one of those areas where it's real easy to decide how we can improve the situation and improve our database. And that's the epistemic uncertainty.

And I just added this one -- where's the mouse? There we go.

When we got into the first consideration of all the omics information and the big data and the data mining and the microarrays and all of these other promising ideas of gathering tons and tons of information, things that when you're working with a toxicologist and you have five animals and three doses, you don't have large datasets in any stretch of the imagination. But if we've got all of this other stuff, we would have large datasets, and we could ask lots and lots of very difficult questions.

And my comment here is beware of going down the rabbit hole, that you think what's going to happen as you zoom in that, you get better and better ideas what's going on, and you

simplify down to the problem until you can just really nail it, what the key event is in this adverse outcome pathway, and you really don't. It's just as difficult as it was before. It's just changed.

Okay, so how do we get -- why isn't it moving on?

Okay, the other way to categorize uncertainty is where the sources are. And so one way is to follow that extrapolation model and say, well, we have the study data, where we start from. We have the goal, the human situation, and then we have the bridge between those two, all the extrapolation methods.

All of these have uncertainties; you heard many of them discussed in the last day and a half in extremely good detail, not the least of which was the homogeneous rodent study and the heterogeneous human population. But some of the things that are now being considered go beyond the simple extrapolation or scaling methods so that the uncertainty factors are no longer isotoxic scaling factors; they're just rough guesses at where we think the range is going to be for the human equivalent dose.

But it's not really a good mechanistic extrapolation in the standard mathematical sense, and we try and incorporate these other concerns such as behavior. I was really pleased to see that show up so often in the studies on tobacco products.

This behavior can dramatically alter your exposure, and then it changes the dataset that you'd like to have for your dose-response assessment. So include that back in that first slide I had on the bridge between the exposure assessment and dose-response assessment, include behavior and the alterations that might happen to influence.

And then the last type of categorizing uncertainty is really in the regulatory process itself.

Uh-oh, I promise I didn't touch anything.

UNIDENTIFIED SPEAKER: Don't roll it.

DR. HERTZBERG: That's it, don't roll it?

(Off microphone comment.)

DR. HERTZBERG: Okay, thanks.

Okay, in the policy perspective, one of the concerns is when is the policy going to change. I remember long ago when I was first with EPA in the 1980s, and a nice quantitative toxicologist from DuPont came up, and he said, you know, it's not the fact that you guys are doing -- overestimating risk that makes it difficult for us. He said we don't have a sense that you're going to let us have the same standard last for 5 years. If we knew that, we could plan, but when it could change next year because of some new study or some new peer review, that makes the whole planning process very difficult to

do. So they're really urging for this, you know, Russian-like 5-year plan put in place, come back in 5 years and rethink it. Superfund ended up doing that, but nobody else as far as I know.

Another one is that the policy can change the actual procedures you're going to do or the perspective, and the example that I gave of the cumulative risk in EPA of going from a source-based assessment to a population-centric assessment is a major change. And now, all of a sudden, instead of having a source as the only exposure, you have to consider the other related exposures that might play a role either in a mixtures context or even just in adding more exposures to the usual population level.

The one about risk addition for carcinogens, there are alternatives, and certainly, dose addition is one of them, and interactions is another one. One of the issues of uncertainty in science is the very poor, often, too often, definitions of the terms that are key to the research publication. We often see many articles showing additive is greater than additive/less than additive, and you say additive, adding what? Sometimes they'll say adding effects, including the famous Presidential Commission on Risk Assessment and Risk Management actually changed their definition in the middle of their

report. Others don't even tell you what they're adding, and if you add response for setting dose, it not only makes a difference quantitatively, but it's a major fundamental change in the concept.

This condition assumes you have toxicologic independence. Dose addition assumes you have toxicologic similarity. If you have similarity, they're not independent. They're jacking right on top of each other and expanding the exposure level. So try and define the terms. Look for those definitions in the articles that you use.

One of the major uncertainties is the defaults because regulatory agencies are often under court order to produce something by a certain time or date, and they have limited resources, so they do the best they can, and you heard several good presentations on examples of this. Many of the defaults, including the dioxin toxicity equivalence factors, were put in place as a temporary interim measure to be replaced by real data soon, when the data become available. They're still there -- what are we now, 30 years later -- as their interim measure. Defaults do not, because of their use, confer credibility, but they certainly give you a sense of at least consistency.

The third one is public perception. We had some great

examples of this earlier of how consumers will change their smoking behavior based upon what they perceive the new product to be or how it works. The filters versus non-filtered, taste changing, so your inhalation rate is different. And certainly, we have the one, the second bullet of "does the public ever follow the advice on the label?" We have too many examples of this in pesticide use where they go out and spray the little buggers until they drop when really just one little puff is all you need. You just have to just hit them once and then wait.

And the last one is very interesting, I think, for communication. Communication came up several times, I think, in the last -- some of the ones today and I think later yesterday. We like to be honest. I think this is one of George Gray's comments perhaps. We like to say what we know, say what we don't know. I remember the first time I was deposed when I was at EPA, and the Department of Justice lawyer said never say, "I don't know." Say, "The state of the science is." And part of the problem is the more you talk about uncertainty, the more chance you have that someone will think you don't know what's going on.

(Off microphone comment.)

DR. HERTZBERG: And so do you admit your uncertainties, or do you try and hide them and give just what you think is a nice

clean example? So that's a dance that you kind of have to do.

The last part of the regulatory process with uncertainty is that all of those three things I just mentioned cross over and influence each other. And so some of the concerns are how much new science do you need before you have a change in the official policy or standards that you're going to follow. An example is the benchmark doses at the no effect level.

How many examples did we have to generate, showing again and again this is a much superior way to do it in general, but not always? It took a long time before it became an official EPA approach, and even then it wasn't well thought out, because people would look at the benchmark dose as the statistical way to handle toxicity, when the counterpart was, as pointed out by several people at EPA, when you have the NOAEL approach, you look at all the studies, you look at all the endpoints, all of the weights of the evidence that you have for different kinds of effects and different kinds of exposures, and then you choose the one that you think is going to be the most health protective. The benchmark dose was too easy to robotize this process, automate it, pick a study, crank out all these numbers and think you now have a good characterization of the uncertainty of the modeling process and the uncertainty in the toxicity, when, in fact, you really do not.

The last two bullets I think are easy to understand, so I won't go into those. We don't know for sure that the male-dominated Congress finally looked at EDCs because there's a male reproductive effect. It just seemed kind of coincidental that they waited so long to do anything.

Now, here's an example of a process to uncertainty, and these are real data from EPA on carbamate pesticides. This is the relative potency approach that Glenn Rice mentioned. This is a true dose addition, and we assume that the component chemicals are like dilutions or concentrations of each other, so once we scale them by potency, we add them up to an equivalent total dose or effective dose for the mixture. And you see the change from 2004 to 2007 and -- I didn't, honest.

(Laughter.)

DR. HERTZBERG: Where did it go? Okay. This is a sensitive mouse.

Okay, you look here and you think well, carbofuran and formetamate must have had some nice new studies that came up, and look how much more potent they are in 2007 than they were in 2004, when, in fact, that's not what has happened. It turns out, with the relative potency approach, you have an index chemical, and you scale everything according to that index chemical, and that's where the change occurred, where BMD10

changed from 0.14 to 0.24.

So less toxic, which made the others seem more toxic by comparison, okay? So the uncertainty wasn't with the carbofuran and formetamate; it was with the index chemical, oxamyl, which calls the question, if that's the most uncertain one, why is it the index chemical? Another time, another discussion.

Here's a conceptual model that's often used in risk assessment for communication, as well as sort of organizing what we're going to do. And what I like about this is that it starts with the pollution sources and then takes them all the way down through fate changes and exposure changes. Some population behavioral approaches can be put in here for the exposure, and then you get target organ tissue doses, a much better idea for exposure than just the source concentrations.

Here's one I'm adding just because of what's going to come after this, just to show you that you can take the exposure levels and take them all the way down to clinical endpoints. So we have here direct pathways, indirect pathways, psychological stress indicators, various risk factors, and then last 2 inches of the slide or so, or 2 inches on my screen, the last third or so are all on these clinical indicators for lung diseases as well as other suggestive factors to worry about.

So that kind of connection can be made.

Unfortunately, we don't have that for all the chemicals, and what I'd like to see and suggest for research is if we could have these benchmark response levels identified for the major health endpoints. When does a change in a certain measure such as arterial sclerosis or hypertension, at what point does it become really important so that the changes of reducing that impact is now medically important, medically meaningful, not just statistically shown?

One thing you see missing here is there's no clue about the uncertainty of this. This and the previous conceptual model have these nice clean boxes with nice clean straight lines connecting them all up. It looks like everything is credible and believable and honest and accurate, and we know that's not the case. We think it's probably a pretty good idea of the linkages, but it's not well done, where it's actually well studied as well as poorly studied.

So here's one I showed my class, and I've done a couple of workshops, and I find it interesting to show -- if I can get the control to show up. Just watch this; this is a commuter train in Germany and noise level, and see what you see. And some of the things that you notice is how protective buildings are, and the small buildings do almost nothing; the large

buildings do pretty good. And if you're standing underneath the bridge, that's one of the best places to be.

And what struck me as interesting was that there are two things missing in the appreciation of this presentation, this animation. The first is sound level is not noise. Noise is sound that's annoying, that you don't like. It gets at you somehow. So the psychological response to the sound is not portrayed here; you don't have a sense of that. And secondly is this took place in like 5 seconds. This is not a chronic exposure condition. It's not even a repeated condition. It might be after -- I mean, maybe Germany has good high-speed trains and it could be happening every hour. But this is not a continuous exposure.

So what you might be able to stand, if you've ever been or have your kids go to a rock concert, you can stand high levels of noise for a short time; it's not disruptive. So be careful when you show some things that you think you're communicating a lot of information and there are concerns, and yet you're not.

And then again, there is no uncertainty indications. And, in fact, comparing this to the conceptual model, you get a sense from some viewers that this is more real or credible because it looks like a real situation. They can put themselves into this picture, into this exposure scenario,

where the conceptual model is just a block; it's a concept. So which one is more factual and more accurate? This is a little misleading because it looks too real.

So here's the future. I was going to talk a lot about ground-truthing with all of the omics data and the big data collections going right now. That was mentioned and discussed before, so I won't say much about that. There was a comment earlier today, I think it was, of becoming more comfortable with uncertainty, and I think, to go along with that, I put in a different context of respect for uncertainty. But uncertainty doesn't mean we are clueless. It means that we have a sense that certain things can be improved by data, certain things cannot. But expressing the information is better than trying to hide it.

One question I was going to have for the previous panel was do you ever have any examples, we could toss it to this panel actually, where the expression of uncertainty is welcomed by the decision maker and not rejected by it. I remember several times at EPA trying to express uncertainty, and the people making the choices for the decisions of cleanup or setting standards said I don't need to know this, just give me your best answer. There has to be a best answer. I'm not sure there's always the best answer, unless you oversimplify what

you think the problem is. So that would be good to know, too, is if you do give a lot of good uncertainty characterization, is it helpful to the final result, which is the intervention or the change that the decision maker has to opt for.

And the last little bullet, understand that your decision makers can suffer information overload. You can't give them everything that you know. You have to filter things out to make it, I think -- as Bill Farland was the one who focused on this -- fit-for-purpose assessments, try and contour all your risk assessments to be useful to the decision that's going to be made and then don't go any further.

And lastly, I came across an article just recently, published in January by an epidemiologist, comparing the toxicology versus epidemiology views of dose response and the chemical exposures, and noted that epidemiology studies or epidemiologists perhaps know too much about humans and how variable they are, so they try and have a very specialized customized approach for every situation that they address.

And he said it would be a lot easier to compare across scenarios, across chemicals, across populations, if we had some core set of approaches, and we could just stick to them and at least use them first and then modify them as necessary. So he is definitely advocating for standard term, standard methods,

at least for the communication and the sharing of information. One of the approaches that I think Dr. Haber showed was the categorical regression for combining severity across different studies. Any kind of meta-analysis, whether you're just going back in time or going across current studies or different scenarios, you really have to have some way to link them to each other, some baseline set of terminology and approaches, or else you really are comparing just apples and oranges, and it doesn't help you.

And this is my final slide, I think. When you get into beyond just the chemicals to include the behavior and the personal characteristics, there are more questions to address, there are more uncertainties to worry about.

And some of these we will not be able to get information on in this country because we have invasion of personal privacy concerns, we have the rejection of governmental bodies coming in to find everything they can about you, we have the potential for hacking of databases. A lot of things are happening in the last 10 years that weren't in existence 20 or 30 years ago in trying to do data gathering. We have cloud sourcing of a lot of information for exposure to personal behaviors to health endpoints. How do we factor that in to our evaluations of these different scenarios and conditions?

And these are some of the people who have helped me out and some key references. Thank you.

(Applause.)

DR. DRESLER: Okay. Dr. Portier, thank you so much for bringing up the last presentation on Uncertainty & Variability: The Rest of the Story, and he's from the American Cancer Society.

DR. PORTIER: Thank you very much. And I want to thank the creators of this workshop for inviting me, and then also at the same time, telling you what were they thinking of, putting a statistician last on the agenda to talk about uncertainty and variability? I mean, it sounds cruel and unusual, right? So I'm going to avoid a lot of that.

The second thing I wanted to say is that, you know, a lot of people have disclosed, so I'm going to disclose I've never worked for a regulatory agency, I've never had to do a risk assessment. What I have done is sat on over 50 expert panels reviewing risk assessment for EPA, NTP, WHO. So I've seen just about everything that's been talked about here over the last 2 days. So what I'm going to -- and then the third thing is, this is the third version of my presentation. I had the first version when they said here's the workshop, and we want you to talk on uncertainty and variability. Then there was the one I

came here with, and then this is the one I redid last night because everything I had prepared had already been discussed yesterday.

So I don't know how this is going to work, but we'll see. So this is the way I looked at this whole workshop. We've had discussion on all of these parts of the hazard and risk process, and kind of linking it in the middle has been these concepts of uncertainty and variability. And my difficulty is that what was left for me to discuss is the red; I guess it's orange on this screen.

You see there's only little pieces around the corners, and these boxes overlap, and they overlap with uncertainty and variability because we've seen these concepts over and over again. We've seen uncertainty and variability discussed multiple times in different scenarios. So there's not -- it wasn't really a lot left for me to kind of elaborate on, add to. So what I'm going to do is kind of go back and take the role of the last speaker and maybe go back and kind of go through this whole 2 days and look at these concepts and how they've been discussed and see if that maybe is where I can add some value to this process.

A number of speakers have talked about definitions. We need to define things, so we need to define uncertainty and

variability. And actually, I noticed Dr. Hertzberg uses uncertainty in general, right, as a general term. So he includes what I'm defining here as variability, which is population heterogeneity, with this concept of uncertainty, which is imperfect knowledge.

And I think in the risk assessment area we use this kind of definition. We want to kind of -- when we're talking variability or variation, we want to talk about heterogeneity and the populations that we're interested in and/or the characteristics of those populations. And that variability is typically quantifiable.

It's not totally knowable, but we have a lot of surveys, we have a lot of census data, we have a lot of sources for understanding things. We know roughly what the distribution is of how many cigarettes a day a smoker smokes; there's pretty good data on that. That's variability. And on the other hand is uncertainty, which is this imperfect knowledge. And in the rest of my talk I have "uncertainty" as lowercase and "UNCERTAINTY" as uppercase. When I have uppercase, it means it might be quantifiable; it's something that we can directly input into a risk assessment. When it's lowercase, it's something that's going to take subjective discussion. It's the kind of thing that experts sitting around a table are going to

decide and/or you're going to discuss with a peer review panel because it's not necessarily easily quantifiable.

And we've been back and forth over all the presentations. We've been talking about some quantifiable uncertainties and some non-quantifiable uncertainties. So my whole talk is when are we talking about which? That, you know, simplifies my whole talk down to that.

So hazard characterization, where we talk about problem formulation, scoping, and framing, there's little case, lowercase uncertainty about the definition of adverse or adverse health effect, you know, so that's not something we can easily quantify, but we've got to define it way up front in the hazard process; otherwise we don't know what our target for our evaluation is. As I thought about the discussion on weight of evidence and scoring of weight of evidence, so yesterday we had some discussion right off the bat about weight of evidence, and today we've had some discussion about how we actually might be able to quantify that and put that into the risk assessment.

And the way I thought about that is that this weight of evidence scoring is a way of quantifying the uncertainty and quality value and fit-for-purpose of available research findings. That's what we're really attempting to do when we go through that, and I think it's great. And I'm upset because I

was hoping to be the first one to use fit-for-purpose, right? I mean, that's a D.C. term these days, but it showed up twice before today. And it showed up yesterday. I was surprised.

What we really -- a big uncertainty in my mind is what's not been studied or measured. So, you know, where are we having to make some big assumptions, you know, where do we have to borrow -- this whole concept of similarity, where do we have to borrow from similar things? That's a big assumption we're making, that this is similar to that, and we've had a lot of discussion on that. But that's a lowercase uncertainty. I'm not sure we're going to be able to actually quantify that decision; that's more of an expert objective decision.

And another one is when is there just not enough information to proceed? You know, it's kind of like you're looking at a body of literature, and you're saying there's not enough here for me to be able to say anything. That's a very difficult decision because you really don't know what the universe looks like. You don't know if this is everything or this is just, you know, 5% and if we wait 5 more years, we're going to have a lot more information. So that's a big uncertainty there.

Uncertain values of data from emerging technologies and tools: Another small uncertainty is what are these emerging

technologies going to provide us? And, you know, we've had a lot of discussion about Tox21, QSARs and how we use them, but there's a lot of big assumptions we're having to make to translate that data to information. And, you know, as Dr. Roberts was saying, it's up and coming, it looks good, but we're going to need to know a lot more before we know what that signal on some kind of screen, what the tox screen looks like in terms of -- or what it's telling us about risk.

Functional pathway models and AOPs, uncertainty as to whether this constituent uses a previously unseen AOP. So the whole concept there is that we have a portfolio of AOPs that we're going to walk into the hazard characterization with, and we're going to look at the available literature and kind of say these are the ones that might be working. But what happens if that portfolio is incomplete? There are mechanisms that we don't know about. Then we tend to -- we may be losing a key hazard because of that.

There was some discussion yesterday about endogenous chemicals with exogenous exposures and this concept of a tipping point. I still see that as a lowercase uncertainty, not quite sure we really finished this conversation. Are there tipping points? Are there levels below which there's really no effect?

I have a hard time -- Dale and I were talking about this at the break, about cancer. I really have a hard time seeing a carcinogen at a level that doesn't have that mutagenic effect. But if there is a tipping point, then it becomes a variation problem because that tipping point may be different for me than for you, all right? So that's a variation issue.

Mixture exposures: And I'm not even sure I want to get into that, but there's a lot of variation and uncertainty in there. There's a lot of variation in composition; that's something that's measurable, we can look at. There's uncertainty in mode of action, in response addition, interaction effects, all of these other things we've been talking about, lowercase. I don't think we're quite to the point yet where you're going to be able to quantify that uncertainty.

Really, one of the things I picked up on is this 10 key characteristics of human carcinogens. I've got to go look at that. Attempt to reduce uncertainty in final characterization of a carcinogen. So they're kind of putting a box around what do we call a carcinogen. So at least now, again, at this hazard characterization phase, we've got a checklist, and we can go does it do this, does it do that? No, no, no, no. It's unlikely to be a carcinogen. Do we need to move much further?

Oh, an aside here. I do want to answer a question I heard yesterday, and I thought it wasn't really well answered, how to identify a similarity mixture when there are thousands of chemicals in the mix. And I think that the answer to the question is not to focus on the thousands. The answer is to focus on, you know, what are we trying to get at? We know that when we look at all of these chemicals, their potency is probably going to be exponential. There will be a small group that have high potency and then everything else, their impact dies off pretty quick.

And yet it may be long-tailed, but you know, it's unlikely that the bulk -- you know, beyond a couple of hundred at most, right? Maybe a hundred even. So why -- you know, don't even think of the hundreds; think back in the fifties and -- I mean not the thousands, think back in the fifties and the hundreds, because I think a few constituents are going to make up the bulk of the mix, and they're likely to drive any risk assessment.

Now, the uncertainty here is how do we get a good ranking of potency? You know, how do we really figure out which of those two, three dozen chemicals that we should be focusing on at this hazard sign.

Exposure route, duration, timing: Exposure data

invariably informs variability. You know, we're looking at product use, demographics, geographics, time. The uncertainty here, and it's unquantified, is all the biases and all the methodologies that we use to estimate exposure. It's going to be very hard to put a number on that uncertainty. But what we're really trying to get at with this exposure data is variability. We have this discussion of breathing zones, secondhand smoke concentration, and they use the phrase within and among or within and between variation, and I wanted you to think about that.

When we talk about "within," we're talking about something, variation within an individual, and then "among" means how I differ from you, right? So it's people to people and within the person. Both of those are variance terms, right?

I vary in what I do in my daily activities, and my daily activity is very different from yours. So both of those are variation terms; that's not uncertainty. We need to explain what a distribution, how that works. And then there's variation in things like the time series obtained for measurement devices, time activity patterns. Now, when we go and start modeling these things, modeling time activity patterns, so not only are we going to have variability in my

model view from your model view, but there's going to be uncertainty because I'm not going to be able to estimate those model parameters exactly. I'm going to have to sample to be able to develop that model and it's -- the uncertainty comes from the sampling distribution of the statistics that we're estimating from those data. So that uncertainty is kind of added on to these two sources of variation.

Smoking use behavior. Again, we talked about within and among, within sessions of one individual, among individuals in a population. Again, this is variation that we're trying to get our hands on. Variation in product content, variation in mouth level exposures. Uncertainty in translating machine smoking yields to humans. Again, because we're going through some kind of model, we're going to have estimate model parameters. It's the uncertainty that we're interested in here.

Population variation: We heard some discussion about metabolic polymorphisms. That's population variation and biomarker response to dose.

Dosimetry models: It gets more complicated, right? So we had among individual variation in some of these factors, things that might be tied to how my body works versus your body works, my age, my disease state. All of those things are going to

factor in to this variation of how we're going to respond to a dose. But typically, there are statistics used to index the distribution of the individual factors. So, you know, when we're talking about variation, we're going to have a mean and a standard deviation or a mean and some other parameter that describes that population variation, and we're going to have to estimate that mean, and there's going to be uncertainty on that.

And when we're interested in both looking at variation and incorporating that uncertainty, that's not additive; it's compounded, in a sense. You end up doing this 2D Monte Carlo or Bayesian estimation. Again, what we're trying to explain is variation, but at the same time we're adding in the uncertainty of the parameters that were used to explain that. So we're kind of adding uncertainty on top of variability.

And one of the statements yesterday, a couple of processes are assumed to be independent and, you know, this is the first time I started thinking, well, there's additional biases that we have to start thinking of, and those biases will come out of -- typically out of assumptions. I mean, we have biases from the data that we collect, and the things we can and can't do when we collect data, and those limitations do cause biases, but you know, if we're careful, typically, that's not the big

problem. A bias, when we make a big assumption like independent, can have big ramifications throughout the whole risk assessment, and so we have to be very careful. Any time I see an assumption, I immediately start thinking, as a statistician, how is this going to bias these parameters moving forward?

We had this discussion yesterday and again today on this role of peak exposures versus average lifetime, and this is an interesting discussion, and I've seen it a number of times in the last 20 years. From a statistician's point of view, there's a big difference between a mean and a maximum.

A mean has nice characteristics, you know. We have this -- you guys all took statistics. The central limit there that tells you, well, when you estimate a mean, its uncertainty has a normal distribution, and so you can use the mean and the standard error to explain that uncertainty in that parameter. When you go to a peak, now you're talking about a maximum, and a maximum is extreme value. It has a nasty distribution. It's got -- you know, it's skewed, it's got a long tail, it's got difficult parameterization, it's got difficult statistical characteristics. Dale's over here shaking his head because he wants me to say, but the area under the curve is the best measure to use, right? That's Dale's answer to everything, the

area under the curve.

So we've had all this discussion about points of departure and estimating. A lot of it is about estimating the point of departure. In my way of thinking, the point of departure is the statistic; we're interested in the uncertainty of that statistic. We've already integrated variability when we get to that point. This is an estimate of the number from which we're going to derive cancer slope factors, inhalation risk, whatever. Everything we talked about this morning kind of jumps off from that POD.

And a lot of the discussion on points of departure is about uncertainty. Sometimes we're going to put variability in there, and that's when we start talking about do I have a cancer dose-response curve that's different than your cancer dose-response curve, all right. And now we're getting into real complicated analyses that, again, is this 2D Monte Carlo/Bayesian type of analysis where we have to be very careful how we put that in. And I think this variation and uncertainty combination shows up in these discussions of duration of exposures, of latency, of recovery times or repair rates. All of these things are the same thing; they're statistics. Uncertainty is what we're interested in, but in some cases we're going to want to conceptualize this, that this

is personal to me, and therefore I have to take into account population variability.

I think we talked about polymorphisms. So susceptibility, and I guess in the discussion on susceptibility is first we got into uncertainty factors, and this is where it gets really complicated because we have uncertainty factors, and then we have uncertainty of the uncertainty factors, right, because you can conceive of an uncertainty factor as a statistic.

You know, it's if I have a PD or a PK model, I can estimate one part of that uncertainty. That is the statistic that comes out of that model; there's going to be some uncertainty associated with that estimate. So I can talk about uncertainty of uncertainty factors. And there was some discussion yesterday about how do we do a better job of setting these.

PBPK models: When we get into PBPK models, you know, they're meant to model some kind of average process, but the way that process works is going to be individual to the person, so there's variability in the realization of that. My transfer rates are going to be different than your transfer rates, so there's variation that we have to take into account. But at the end of the day, we also have to estimate all of these parameters to make the model work, and there's uncertainty

there. And there's a lot of two-dimensional Monte Carlo/Bayesian PBPK modeling that goes on.

And, you know, I just wanted to make sure everybody -- we really haven't talked about 2D Monte Carlo or this kind of Bayesian estimation. So this is the 1-minute seminar on it, right?

So we take *y* sum value or measure of characteristic of an individual in the population of interest. To say *y* varies between individuals is equivalent to assigning to that characteristic a population distribution, some probability distribution says the likelihood you're falling on that distribution. Think body weight, something simple, right? So that distribution is key to a couple of unknown parameters, that delta-1 and delta-2. That probability of distribution describes the population variability, all right; it's the statistical way we tell you what this population looks like. In our way of thinking, that F(delta-1, delta-2) is population in my mind, right? They're one and the same.

But if these distribution parameters are unknown and we have to estimate them, then at the end of the day we're going to put estimates in there, and there's going to be uncertainty. So we might assign, in this case, a multivariate probability of distribution to these two parameters that describes its

uncertainty. And then we're going to use what's called a compound distribution.

We're going to look at that population distribution and integrate through the uncertainty distribution, and that's going to produce -- we're not just summing variability here. We're now going to have a population distribution that's going to -- it's going to be spread more, it's going to have more variability, but it will now have incorporated that uncertainty in it.

I'm running out of time. I want to just skip this one, anyway.

Cancer risk assessment: You know, we look at it as exposure by the cancer slope factor, the cancer slope factor jumps off the point of departure, so we've got that uncertainty that we talked about. When we model this, when we model the dose-response equation in order to get to the point of departure, that point of departure becomes a statistic, and we have an uncertainty with it.

And the way we typically account for that uncertainty is to look at a lower confidence bound. So, you know, that's kind of a simple way where we've incorporated uncertainty. That's not variability. We haven't assumed that you and I have a different point of departure. We assumed that there's one that

kind of represents the whole population.

One of the things that I'm seeing more of recently that we haven't talked about here is this concept of model averaging. So in the BMD modeling, you're assuming a shape for this dose-response curve, and that shape helps you estimate that point of departure. Well, what happens if you go into the EPA's standard tool and you fit nine different models, and they all fit approximately the same? Okay. Now, they're going to fit the same where there's data, but they're going to have very different spreads when you extrapolate below. So if your data is not at -- I mean, if you haven't observed something at a 10% response level, you're going to be extrapolating down. These models have different values. What we're finding is that, well, you know, it's cheap to fit those models. Fit them and average them, okay, and you get a better statistic with better uncertainty properties.

By that, I mean the uncertainty distribution is tighter. So, you know, this is kind of one of the ways we're seeing kind of modern simple computational statistic, not near as complicated as a PBPK. This is just get some curves, get an estimate, average them, wide spread.

Yeah, risk variation. So now we're down to risk, and when you think about it, at the end of the day, if we're taking into

account variation in exposure and -- so we can take into account variation in exposure, uncertainty in that point of departure which leads to uncertainty in, say, the cancer slope factor or the inhalation, you're going to end up with risk that varies, right? You're going to get a distribution of risk values, and that's what we're trying to get at, that tells us something about our expectation of seeing adverse events in this population, right?

When we add in the uncertainty from the point of departure and use something like the lower 95% confidence limit, then we're actually taking into account both uncertainty and variation. And I'm at the red, and I told them I wouldn't get a red.

These are some kind of various statements that have come, that I picked up on; they're less important. I just want to kind of conclude with bringing us back to, well, you know, why do we want to even put uncertainty and variability in a risk assessment? You know, I keep coming back to this.

This whole process is aimed toward risk characterization, which is aimed toward management, right? This is a very practical, pragmatic process that we engage in. So if incorporating uncertainty and variability into this process doesn't increase the value of the risk characterization, we've

wasted our time.

You know, it's nice to think about the statistics, but it better add value. We had this conversation yesterday, like when do I know to stop modeling? From my point of view, I stop modeling when it no longer adds value. You know, if I'm going to increase understanding by one-half of one-tenth of 1%, why go to that effort? It just adds complexity without a return, without benefit. So variation ensures we don't over and underestimate risk. Uncertainty ensures we don't, or we take into account data limitations. Better risk characterization, better decisions, better regulation.

And that's kind of why I continue to volunteer for all of these peer review panels. Not only is it fun, but I have the feeling that bringing this kind of uncertainty and variability and understanding to play in risk characterization does produce a better product, which does produce a better decision on the part of the agency.

Thank you.

(Applause.)

DR. DRESLER: Thank you.

So, again, questions to the white cards. But I have one that was from the last session that I said I would go ahead and ask here. So how is uncertainty and variability considered

with current methods for point estimates, non-probabilistic risk values? So how is uncertainty and variability considered with current methods for point estimates, non-probabilistic risk values?

DR. PORTIER: Badly.

(Laughter.)

DR. DRESLER: Oh, okay.

DR. PORTIER: So I think what they're referring to is probably just incorporating uncertainty in a discussion, right? I mean, that's the traditional thing. Well, I have a statistic, and I know confidence intervals on the statistic, so I'm not really going to this full probabilistic approach where I'm integrating variability into this whole process. I've just estimated a value in from the data, whatever the fit is. Here's the uncertainty in that estimate. So, you know, at best you're talking about a point estimate with a confidence interval around it, you know, and that's very traditional. I mean, that's what we did 20 years ago when I started this process.

DR. DRESLER: Okay.

DR. PORTIER: You're on, Rick.

DR. HERTZBERG: Okay. I think there are two areas that may be worth commenting on, and one certainly is the process

that when you have point estimates such as the reference dose, that one of the concerns people have when they publish a paper using a certain reference dose, the good ones that include uncertainty into their discussion mention when that reference dose was created and what were the underlying datasets. And what was kind of the norm for the studies done at the time when the critical study was performed, that was eventually selected, because sometimes you had differences in technology and differences in certain accuracy of measurement, and other times you just have a gap of 30 years or so where no one cares about that chemical and so no one studies it. So the information you might have would be a multivariate about newer chemicals or chemicals that are in the public eye.

For the particular one of interest, it doesn't exist at all; there are no baseline studies, no key fundamental studies at all. And so the best you can do is just acknowledge the fact that this is based upon old information, not necessarily bad information, but old information that's usually imperfect and partial, and that's the best you can do.

DR. DRESLER: Dr. Hattis? No? Okay.

Okay, so let's see. The extrapolation of effects in animal studies to human effects can be particularly challenging in tobacco-related research due to physiologic and behavioral

differences. What methods have been used to determine whether epidemiologic or toxicologic information are scientifically valid and whether extrapolations from animal data could address those data gaps? So let me do that again. So the extrapolation of effects in animal studies to human effects can be particularly challenging in tobacco-related research due to physiologic and behavioral differences. What methods have been used to determine whether epidemiologic or toxicologic information are scientifically valid and whether extrapolations from animal data could address those data gaps?

DR. PORTIER: So let's separate the scientifically valid from the extrapolation question. I mean, a big part of peer review, a big part of the quality review and weight of evidence is really looking at not just the value of the studies, but the quality of the studies, right? I mean, EPA calls it a quality review.

They're looking at the design, and what makes a good scientific study valid is a good, well-documented design, right, that has good properties that don't have obvious biases, and where the research team that's done the work and has been honest in writing up the results, right? One of the biggest problems I've noticed as I've reviewed literature is I see a beautiful study, a good experimental design, a clear

description of the results, and then a discussion that goes off into another galaxy, you know? A good study with a good team, you know, they keep their conclusions close to their data; they don't extrapolate too far out. To me, that's my definition of a good, valid scientific study.

And I think the same thing for epi studies. You now, it's the same way. I mean, they're more complicated, there's more opportunity for bias in an epidemiology study, and each of the designs, you know, cohort, longitudinal, cross-sectional, they all have good and bad characteristics, but most of us who deal with that stuff understand those characteristics.

The fact that you do one kind of study over another doesn't change that study's validity. It just tells us I have to look at the results a certain way, and I can be -- I'll have to be careful how I interpret those results, but it doesn't change validity. Now, the second part of the question was extrapolation, right?

DR. DRESLER: Yes. And whether extrapolations from animal data could address those data gaps.

DR. PORTIER: Yeah. So, you know, the mathematician in me says that, you know, if there's no data, extrapolation assumes I'm going beyond my data, right? I mean, you have your data. If you're in between the data, you interpolate; the minute you

step outside your data, you're extrapolating. You can only extrapolate with a model. Some models are simple, some models are complicated. These PBPK models that Dale loves and we've discussed in many panels is just another very complicated model. But all extrapolation requires a model, if it's a simple regression or a complicated mechanistic construct.

So the -- and what was -- the question was how do we do it? We do it through a model, and it jumps off from good data. So if you get good valid data and you have a good grounded conceptual model, you probably have a pretty good extrapolation.

The other thing that statistics tell us is that the further we get from the data, the more our uncertainty, and it's actually hyperbolic, right? So you can go two, three units, but the minute you go in order of magnitude, you probably have 10 orders of magnitude of uncertainty, and now that extrapolation is essentially useless because the uncertainty swamps the value of that number. So you can only extrapolate so far. And this is the big part of this low-dose extrapolation argument, you know; we're doing animal studies way up here, human exposures way down here. What's the model that takes me down there? And every model that we apply, it's so far from the data to the exposure we're interested in that

our uncertainties get huge, and then we're left with a lot of scientific conversation and a lot of panels meeting and talking about it, right?

DR. HATTIS: Well, the validity of the projection of observations from animals to people depends upon whether the starting point and the ending point, the animal system and the human system, share mechanistic characteristics that allow you to reasonably predict people or whether, in fact, there are such profound differences that you're unlikely to be capturing what matters to determine the outcome that you're trying to predict. And so that's basically the issue of "validity."

If I don't have -- if I'm projecting from a non-pregnant animal to a pregnant person, then I have to worry about the pregnant person has an especially sensitive fetus inside. Am I really appropriately representing that with the observations in the animals? On the other hand, you know, if I have a pregnant animal and a pregnant person, then I have to say, okay, is my exposure on Day 6 of the animal's gestation, how does that -- is that appropriate for the 21st day of the human gestation? I mean, what is the right analogy to make? And I have to worry about those kinds of things in constructing both the best system to make the projection and to see whether I've got effects that are reasonably predictable from the data that I

have.

DR. DRESLER: I wonder if this next question is going to kind of follow along. Did you want to go ahead, sir?

DR. HERTZBERG: Yeah, just let me add a couple things. First, I think the question was partly, maybe, what kind of information we would like to have to feel that an extrapolation approach is valid or somewhat accurate.

And certainly one of the areas is just gather some data, no matter how imperfect or partial the data might be. If it's for the situation of concern, then you at least have some way to start ground-truthing whatever your extrapolation is. Some of these approaches are certainly biomarkers, exposure biomarkers and effect biomarkers. If you can show the same kinds of sequence of biomarkers as relevant for both the animal and the humans, then you get a better sense that perhaps the key steps in the adverse outcome pathways are similar as well.

The next comment is just that there's also extrapolation that's really a misnomer for some species changes that's really just scaling, and the scaling is usually not model based. It's just either based on the decimal system, so we use a factor of 10, or it's based upon the history where certain agencies like FDA and EPA have been using 10, 100, and 1,000 for so long without great public disasters ensuing, so it must be sort of

in the ballpark. There is some good validity in the risk assessment process of seeing what seems to work, even if it's not scientifically very accurate or very mechanistically based.

And then the last thing was the comment on the metric, that the metric should be used if it has some -- or preferred if it has a biological basis, not just because everybody else does it. So the good examples that we had of oral exposure is not the same as drinking water exposure. I thought oral was, you know, gavage or drinking water, and there's not much else, and now we have, well, it just depends where in the mouth you're talking about and for how long, what the duration is. That kind of specificity needs to be reflected somehow in whatever the dose metric is that you're using. If you're not using that kind of metric for the animals as well as the humans, then you're missing something that's very important in this whole dose-response characterization.

DR. DRESLER: This follows up on this a little bit, what you brought up. As the diseases due to tobacco use may take years to manifest and discover, the understanding of the time to disease onset will be valuable information in risk assessment. What kinds of studies may be conducted to obtain biomarker and disease information for users of tobacco products?

DR. HATTIS: Well, there's a lot of hope for much better use of biomarkers for risk assessment as the -- our scientific techniques and understanding. One that is important, I think, or going to be important for carcinogenesis by genetic mechanisms is the use of information on mutation frequencies from exposed tissues to predict mutations that occur along known pathways to cancer. So if you have, you know, four, five, or six specific mutations that you know are associated with enhanced carcinogenesis, and you observe those in smokers versus nonsmokers or people with one kind of characteristic versus another, then you have a good chance, even before the cancer develops, you have a good chance of quantifying the enhancement of that kind of the cancer process, even in people. And if you could then relate that to analogous observations in animals, you have a much more powerful tool for making predications even, you know, at a much shorter time frame than is necessary for the full cancers to become manifested.

So I think there's a lot of hope for the uses of biomarkers not only for detecting, for better quantifying dosimetry, as hemoglobin adducts and things of that sort, but also for detecting changes that you know are already on the pathway to end effects.

DR. PORTIER: So I'm not an epidemiologist, but I

occasionally have to channel my colleagues, and they would say, well, the kind of study we're talking about is the American Cancer Society Cancer Prevention Study-3, which we initiated a few years ago, recruiting 350,000 people and getting blood samples from all 350,000 people, and then we're planning to follow them for 40 years doing the biomarker estimates. A lot of that work is being done now, but we're saving a lot of that data for some future point when some different AOP needs to be looked at and a different biomarker becomes, you know -- and as we know, that kind of research is incredibly expensive, incredibly time consuming. We estimated in our CPS-2, which was a longitudinal study of a little over a million people, lasting 40 years, costs probably over $60 million in the lifetime of that study. You know, that's gold standard data.

I mean, at the end of 40 years, we'll have an answer to that question. Unfortunately, I don't think we want to wait around for 40 years for that data to come in. So we're spending a lot of time talking about what can we do now? You know, do we have to wait for all 350,000 to die before we have answers to some of these questions?

And the thing about tobacco, you know, is we know something about the latency of tobacco. You know, that's what, 30 -- 25 to 35 years. So some of the people in the study have

been smoking for 20 years before they joined us. They just didn't have to have cancer on the day they filled out the questionnaire and gave us a blood sample. So we already know that a number of them are going to start to, unfortunately, come down with cancer in the near future, but we will have been able to establish the biomarker data, we'll be able create the controls and the comparison populations to be able to establish what is signal and what's not signal in these kind of things. But it's still going to take time.

DR. HERTZBERG: Yeah, I'd just add to that, that I think a couple things that I noticed from talks this morning and yesterday is that there are certain chemicals where the effect is predominantly because of peak exposure versus long-term, time-weighted average exposure. So if you have a long time waiting between whatever the exposure period is and this effect of the long latency period or latency time, the peak exposure may have come and gone, and you may not have it even in your dataset that the peak exposure occurred.

So there could be a case where you expect a case to show up later on, but it's an exposure that was shorter term and was just totally missed. That's one of the issues with a lot of the NHANES data. And then the other one was that some of the effects are reversible, and so if you have biomarkers of

effect, not just exposure, it would be good to have measurements long before the 20 years or 30 years latency period. But just be aware that some of those effects are going to correct, and so they may not show up at the 15-year point. They might be important later in the whole etiology of the disease.

DR. DRESLER: The risk or harm of tobacco use may induce multiple health outcomes. What approaches may be used to quantify single health outcomes among possible multiple health outcomes that may occur concurrently? So what approaches could be used to quantify the risk from aggregate health outcomes?

So when you use tobacco there's often multiple health outcomes, and so what approaches may be used to quantify single health outcomes among possible multiple health outcomes that may occur concurrently? What approaches could you use to get that aggregate health outcome?

Now, you'd think there would be easy questions at the end of this, I mean, really. How about if I -- let me say it again because it is tricky. So using tobacco causes multiple health outcomes. What approaches could be used to quantify single health outcomes amongst the many that may occur or do occur concurrently? So what approaches could be used to quantify aggregate? And you can imagine that would happen in a tobacco

product. You asked, okay, I wanted to hear about COPD, which we haven't talked about much here. We talked about cancer and noncancer, but COPD, okay? But COPD is going to happen somewhere in that inflammatory pathway of cancer. So how are you going to look at that, but look at it aggregately? I don't know the answer. I'm asking you guys.

DR. PORTIER: Yeah, I don't know the answer either but, you know, what confounds the answer, and again, I don't know, is if cancer is in that mix, you know, especially the cancers that we often associate with tobacco use, you know, people die pretty quick. So even if they have the markers and the conditions for COPD and heart disease, cancer eliminates that.

So there's a competing risk going on here that eliminates those people from the population that would help us look at this joint risk model, which is really what you're trying to get at here. I want to be able to look at the joint risk and then marginalize, margin out COPD and say, well, if I didn't have cancer and if I didn't have heart disease, what would my distribution of COPD look like as an effect from tobacco use? And you have to be able to look at the joint, and with cancer in that mix as a censoring agent, it makes it almost impossible to even think of how I would collect data. If cancer weren't in the mix, again, I could do a cohort epi study and follow

people and see, well, when does this show up and when does that show up? And eventually I'd build up enough data to be able to build that multi-disease model.

DR. DRESLER: Okay, let me revise it. So in the U.S., most people get lung cancer somewhere around the age of 70 now, okay, so it's around 70. So let's take a 42-year-old that's been smoking; average age of initiation is around 13-14 years, okay? So they've got several decades they've been smoking, and so we don't have a cancer risk per se in there, but we've got cardiovascular, but your single health outcome is diabetes. So take the cancer one out of it because we're talking about 42-year-olds. Does that help at all?

DR. HATTIS: Yeah, you've got to continually do life table-type analyses that account -- that represent in your spreadsheet or other device the censoring, the progressive censoring of the population, because as I proceed in my analysis of FEV_1 distributions yesterday, I've got to -- and I didn't do that yet, in that analysis -- I've got to factor in the understanding that the people in my higher exposure groups are more censored because some of them have died from heart attacks and other things than the people in lower exposure groups. The challenge is to make a dynamic model that appropriately reflects the increasing censoring of the

population with progressively larger cumulative exposures. And it can be done, but it requires thought.

DR. DRESLER: And I'm seeing -- so okay. Anything else?

DR. PORTIER: Yeah, I was going to say it hasn't been done. I mean, I would've thought with some of these very large epi studies that follow not only smoking habits but disease outcomes, you'd have that joint occurrence data. I mean what, the Nurses' Health Study or something like this, where we've been following, you know, 60-, 70,000 people for 25 or 30 years.

You know, if what you're really saying is well, that cancer outcome is really occurring after that heart disease initially or that COPD event, you should see those events occur and be able to look at their time frame. Now, what Dale was mentioning is still a model to tease out those effects, right, trying to figure out is -- you know, do these things work together. If COPD actually exacerbates heart disease, right, then you know, you should be able to see from the COPD event, it's a shorter time, the heart disease, than somebody who didn't get COPD. But there are models for looking at that. So I would think that could be done. There's probably data available.

DR. HATTIS: Data are likely available. It probably

hasn't been done probably because of loss via science orientations that --

(Off microphone comment.)

DR. HATTIS: Yeah. I mean, basically, epidemiologists tend to have a more positivist philosophy of science orientation, which means that they may not take as much advantage of opportunities to do theorizing as a risk assessor would. I mean, a risk assessor learns that the added value that he or she can generate involves theorizing. And so different specialties have different orientations toward analysis and can produce different kinds of information.

DR. DRESLER: Okay, one more question. Dr. Hertzberg, did you want to say anything?

(Off microphone response.)

DR. DRESLER: Okay. Most reference values are deterministic. When performing a probabilistic analysis, what approaches may be used to capture the variability in that same -- that surrounds that value? So most reference values are deterministic. When performing a probabilistic analysis, what approaches may be used to capture the variability that surrounds that value?

DR. HATTIS: Well, I think, basically, the idea is that you need to replace the uncertainty factors with distributions

that reflect separately both uncertainty and variability, and you do that by assembling as many cases as possible where the concern represented by a particular uncertainty factor is captured.

So I basically, for example, have assembled a large amount of data on how much pharmacokinetic variability there is from studies in drugs and things like that, and how much pharmacodynamic variability there is from things like I showed you, and basically how many people get a particular amount of FEV_1 change acutely in response to a particular ozone exposure or exposure to some drug in a clinical setting where I know what the dose was. So this exercise in replacing the safety factors or the uncertainty factors with distributions is an enterprise that will take quite a while and quite a bit of research, but we've done some of it already, and the opportunity, I think, is to use what we have and try to shed additional light for both decision makers and their constituents about what real likely variability is and how it changes risks, and what a fair description of the uncertainties are in particular risk estimates.

DR. HERTZBERG: Well, I think the problem is a bit more complicated than -- or at least the way I interpret the question is a bit more complicated than just replacing the

uncertainty factors. I mean, ideally, some of these uncertainty factors are actually isotoxic scaling factors, and so they should be replaced by biologically based models, whether pharmacokinetic or pharmacodynamic models.

But to get rid of that kind of simple single constant as representing all of the dose-response curve change from one condition to the other, we know a lot more about that for several chemicals, and we should start doing that to just get rid of the factor, whether it's distribution or deterministic, and replace it with something that's really biologically based. The second one is -- there's an uncertainty factor for database deficiencies, and this is where I jumped up almost in my chair seeing George Gray's database on pesticides, that there's an opportunity there to go kind of back in time and then pretend you have certain information and not other information and see what difference it would make on your approach to calculating a reference dose using some standard framework or procedure to do so. What would change if you added the other information?

There was a study done by Mike Dourson many, many years ago of whether reproductive data was ever valuable for improving the reference dose, and it was almost never influential on making a significant change in the reference dose. So it's valuable in the sense that it's one of those

public concerns of reproductive success, but in terms of actually the regulatory use and the reference dose, it wasn't very useful.

The other uncertainty, I guess, in the reference dose is one that I mentioned of the process, that you could say how likely is it that someone's going to do something different or sufficiently different that the reference dose will be -- by dictate or policy, will be changed, will be forced to be changed. And that's kind of like the database deficiencies of are they sufficient that a new study will be really important enough to force a quantitative change in what's listed as the official value. You could do, again, retrospective evaluation of existing datasets and just see how often that they are updated. When they're updated, is it ever a major effort or a major change, or is it just something minor to reflect the current studies?

The disadvantage with using Dr. Gray's database on pesticides is that there are other chemicals that are very, very different than pesticides. VOCs are different, EDCs are different, metals are different, except those metals that are pesticides, I guess.

So you have to be careful that you're not just inferring too much from a very particular dataset. So I think there are

a lot of things that can be done. Someone just has to go do it. There's not a lot of incentive to do it, because if you --

DR. PORTIER: Actually, it's --

DR. HERTZBERG: -- have a reference dose and --

DR. PORTIER: -- already been done.

DR. HERTZBERG: Well, some of it has been done, yeah. But if you have a reference dose that's on the books or an MRL on the books and people are using it and not complaining, then there are a lot of other things to worry about. There are, you know, another 58,900 chemicals or whatever the total is.

DR. DRESLER: Dr. Hattis, were you -- okay.

DR. PORTIER: I was going to say my flip answer was you need to go back and read the minutes of these last 2 days, because all the methodology to do that has really been discussed. And then my second answer is I have seen it done before really well, and you talked about the cost. Well, the report was 1,000 pages, the supplement was 1,000 pages, and the presentation before the panel took 3 days, just to replace that 3 by 3 by 10 by 10 with real mechanistic, good, data-driven estimates.

It's not easy to do, it took a lot of good technical staff a long time to put that together, and not easy to understand. And at the end of the day, it wasn't that different than 3 by 3

by 10 by 10 or whatever it was that they used the uncertainty factors. So at the end of the day, you wonder if the models were picked because they got answers like we like, or whether it just was, you know, it just fell out that way. I chose to believe it fell out that way because they presented it very -- this is EPA -- very straightforward, and it was great, you know, did all the nice extrapolation from mouse to rat, rat to human, human to sensitive human, all in a model framework. It was beautiful. It took forever to read, a thousand pages.

DR. DRESLER: Dr. Hattis, were you going to add something?

DR. HATTIS: Yeah, in the specific case of the database uncertainty factor, we did an analysis of this kind using the pesticides as the datasets, about 60 out of them, and we basically got an answer that about 80% of the time, it didn't matter, but in the remaining 20% of the time, it made something like, if I remember correctly, we developed a distribution for the amounts of -- for the residual times that it did matter.

In that case, it came out to be much less than a factor of 10 as being a sensible value to characterize it. But basically, it's no -- once you have that distribution, you can easily represent it in a spreadsheet-type uncertainty analysis, and so it's -- there are better -- there are better ways of doing this, you know.

DR. DRESLER: What a set of last words, so --

(Laughter.)

DR. PORTIER: Well, my answer, there may be ways of doing it, but there might not be time-valued ways of doing it. I mean --

DR. HATTIS: Well, has anyone ever done a cost-benefit analysis on cost-benefit analyses?

(Laughter.)

DR. HATTIS: You know.

DR. PORTIER: You're speaking to an evaluator who runs an evaluation group, and we evaluated ourselves last year.

DR. HATTIS: Right, yes. Anyway, I think --

DR. PORTIER: Totally unbiased.

DR. HATTIS: I think for last words, we need to understand what our contributions can be to the policy-relevant choices that are made by decision makers and their constituents, okay. So essentially, it's important, the idea of variability is relevant to the social policy issue of fairness of the distribution of risks that are assumed or imposed by a diverse population. The issue of uncertainty is highly relevant to the issue of what the tradeoff is between efforts to reduce exposures and efforts to have a good functioning economy. And so one of the -- so I think it's important that we keep in

mind, in our discussions for decision makers, the importance of demystifying the processes of creating these estimates and making it sufficiently clear how they should be able to use these to address these kinds of social policy questions that they are concerned with.

DR. HERTZBERG: I'll just add one recommendation for CTP and their efforts moving forward, and it kind of hinges on a sort of frustration I got from Ken's discussion of his example. And I think that when you try and do the right thing, the right science in risk assessment with reference doses or cancer estimates, whatever, that often it's a very customized approach, and you replace an uncertainty factor with something that's very specific to the data that you have and the information that you have, the number of supporters you have for your position and your model.

And I think that for CTP to move forward on this, it would be very valuable to have a framework, to have something that's peer reviewed, publicly vetted a little bit, so you can have something to fall back on as supported by the scientific community. Then when you have uncertainties that you can address or not address, you have some consistency of approach. You could look across your experiences and see what seemed to work and what did not and then tweak your framework and make it

a little bit more efficient. But going on where every single issue is always a customized approach, I think, is a very inefficient way and not likely to work very well. So I advocate developing a framework.

DR. DRESLER: Okay.

DR. PORTIER: I guess my last recommendation would be don't shortchange risk communication. Sitting in a nonprofit, a large nonprofit like I do, I get those questions about safety of products that we get to our national call center. So they'll call our 800 number and say is this cosmetic safe for me to use; you know, I spend hours trying to answer that in common language, and the only experience in risk communication I've had is a workshop that Steve Roberts set up years and years and years ago that I learned a whole lot from.

But boy, you know, it is really important that the population not only -- I don't think they really need to understand the full process, but they need to feel like you're honestly communicating with them on this concept of safety of which they don't care about the nuances that we have about safe, you know? They just want, kind of, a black and white is it okay for me to use this for the next year, and then I'm changing my product to something else, you know. So a lot of their concerns are not long-term; their concerns are acute, and

yet we're dealing a lot of time with long term, you know, is this going to cause cancer? No, but it's going to make your skin break out pretty bad.

DR. DRESLER: Okay. Well, thank you very much for both the presentations and another great panel discussion. So thank you very much and your words of wisdom, so much appreciated.

(Applause.)

DR. DRESLER: And thank you to everyone who stayed in there and participated both days and hung out through the end of today. Thank you very much for your attention during this, and it was very, very helpful. So thank you all.

(Whereupon, at 4:13 p.m. the meeting was concluded.)

C E R T I F I C A T E

This is to certify that the attached proceedings in the matter of:

RISK ASSESSMENT OF TOBACCO PRODUCTS: A PUBLIC WORKSHOP

November 16, 2016

Hyattsville, Maryland

were held as herein appears, and that this is the original transcription thereof for the files of the Food and Drug

Administration, Center for Tobacco Products.

_____

ED SCHWEITZER

Official Reporter