UNITED STATES OF AMERICA

DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

+ + +

CENTER FOR TOBACCO PRODUCTS

+ + +

BIOMARKERS OF POTENTIAL HARM: A PUBLIC WORKSHOP

+ + +

April 5, 2016
8:30 a.m.

FDA White Oak Conference Center
Building 31, Room 1503
10903 New Hampshire Avenue
Silver Spring, MD 20993

FDA:

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

This transcript has not been edited or corrected, but appears as received from the commercial transcribing service.

Professional Video Associates, Inc.
2515 Saint George Way
Brookeville, MD 20833
301-924-1556

PUBLIC COMMENT SESSION

PAUL NELSON, Ph.D.
CORESTA Biomarkers Subgroup

KERRY S. LANE, M.D.
Palm Beach Autism Institute

SESSION 4: CANCER

STEPHEN S. HECHT, Ph.D.
Masonic Cancer Center
University of Minnesota

MEREDITH SHIELS, Ph.D.
NIH National Cancer Institute

DOUGLAS A. BELL, Ph.D.
NIH National Institute of Environmental Health Sciences

YORK E. MILLER, M.D.
University of Colorado School of Medicine
Denver Veterans Affairs Medical Center

AVRUM SPIRA, M.D., M.Sc.
Boston University School of Medicine

SESSION 5: NEW AREAS OF RESEARCH

D. GAYLE DeBORD, Ph.D.
CDC National Institute for Occupational Safety and Health

STEPHEN EDWARDS, Ph.D.
U.S. Environmental Protection Agency

CHRISTOPHER PROCTOR, B.Sc. (Hon), Ph.D.
British American Tobacco

EMILY VOGTMANN, Ph.D., M.P.H.
NIH National Cancer Institute

G. L. PRASAD, Ph.D.
RAI Services Company

SHASHI AMUR, Ph.D.
FDA Center for Drug Evaluation and Research

ALSO PARTICIPATING

KIMBERLY FROST-PINEDA, Ph.D., M.P.H., CPH
Altria Client Services

ROBERT TARRAN, Ph.D.
University of North Carolina at Chapel Hill
The UNC School of Medicine TCORS Program

DOROTHY K. HATSUKAMI, Ph.D.
University of Minnesota

TAPAS SENGUPTA
RAI Services Company

MOHAMADI SARKAR, M.Pharm., Ph.D.
Altria Client Services

BRUCE CLARK, Ph.D.
Philip Morris International

INDEX

INDEX

M E E T I N G

(8:30 a.m.)

DR. DRESLER: So welcome back to the second day. Personally, I found yesterday quite fascinating, interesting, and educational, so I'm very much looking forward to today. We'll start today at 8:30, and we should finish up around 4:00, perhaps a little bit earlier than that.

The first thing that we're going to do today is have a Public Comment Session. As you may or may not know, all of our workshops, our public workshops do allow for a period of time for the public to submit a request to speak, and then there's the open microphone, basically for the public to come speak versus the invited speakers, for the others.

Just to follow up, if you weren't here yesterday, which I think I'm recognizing most of the faces, the bathroom is down around the corner. Make sure at the breakfast break that you order your sandwich. I'm not sure the line will be so long, but if you ordered a lunch, you will have it ready for you at the lunch break.

So without further ado, let's go ahead and start the Public Comment Session. And I know that we have two speakers that have signed in. So the first the speaker, please, Paul

Nelson from CORESTA Biomarkers Sub-workgroup -- Subgroup. And I know you know how to use those.

DR. NELSON: Yeah. Thank you. I'm Paul Nelson. I'm the Director of Clinical Studies at RAI Services Company, but I'm speaking to you here today on behalf of the CORESTA Biomarkers Subgroup. And really what I'm here to do is encourage participation from those here in the audience, those watching on the Internet, encouraging people to participate in the subgroup because what the subgroup does is directly relevant to what this workshop is about.

Two questions I can imagine folks might generally have are what is CORESTA, or isn't that an industry trade group? So I'd like to take a few moments just to talk a little bit about what CORESTA is.

CORESTA comes from the French name for the organization. It's based in Paris. I am not going to butcher the French, but it translates roughly as Cooperation Center for Scientific Research Relative to Tobacco.

The vision of CORESTA, as here, is to be recognized by its members and relevant external bodies as an authoritative source of publicly available credible science and best practices related to tobacco and its derived products. CORESTA is not

just an industry organization, although the tobacco industry does participate heavily within CORESTA. It is an organization that's made up of industry, academia, and a wide variety of suppliers and other interested parties that all come together to discuss tobacco science from seed to a biomarker measured in a tobacco user's blood or urine.

Overall, the purpose of CORESTA is to encourage international cooperation and to actively work on tobacco-related areas of research. To do that, tobacco has four study groups, two -- it says it's two plus two on this slide, or two by two, ag-phyto and smoke-techno. The ag-phyto is agronomy and leaf integrity, phytopathology and genetics, not that relevant to this group here today. On the other side is smoke science and product technology, and within the smoke science study group, there's the biomarker subgroup.

So the biomarker subgroup meets twice a year, typically in May and October. It will be meeting in Paris, May 11th of this year, and October it will be meeting in association with the CORESTA congress in Berlin. The convener of the subgroup is Dr. G. L. Prasad -- you'll be hearing him speak a bit later this afternoon -- from RAI Services. The secretary is Kirk Newland from Celerion, again, a non-industry participant in the

subgroup.

At each of the subgroup meetings, we typically have up to around 30 delegates from industry, bioanalytical labs, and other organizations to participate in the discussions that take place.

The subgroup has three objectives or kind of three reasons for being. One is the review of knowledge of smoking-related biomarkers of exposure and effect and document those in meeting minutes, CORESTA reports, or other scientific publications. There are a few examples that are up here, the sorts of things that have been discussed within the subgroup meetings. You'll see that most of those we deal with are relevant to biomarkers of effect.

Additionally, the subgroup undertakes ring trials and proficiency tests for selected biomarkers, as agreed by the SC. That's the scientific commission. It's the body that oversees each of the subgroups. In the past, it has been involved in doing nicotine and metabolite ring trials, proficiency testing for 3-HPMA -- that's a metabolite of acrolein -- and currently is about to get under way a look at NNAL, a metabolite of NNK, an inter-lab method comparison. And for this, although it's a biomarker of exposure, I would very much like to encourage any

interested people that might want to participate in this ring trial to get in contact with Dr. Prasad. The ring trial is still open. We can get more people, more labs can participate in that. The plan is to get that under way in the summer of this year.

And the third objective is to source and develop reference materials to support biomarker analysis. In the past, we found that some suppliers of reference material gave very inconsistent reference material. That really hampers the ability to measure biomarkers. And so one of the things that the group does is get there to discuss whether we need a guideline on standards for reference material and could that be developed and presented to CORESTA.

Going back to the previous objective, a thing I forgot to mention, a lot of what you saw, it was around biomarkers of exposure. That doesn't preclude the group or the subgroup from looking at biomarkers of effect. There are not a whole lot of people working together on the same method in that area, but that is something that the subgroup could undertake, is looking at biomarkers of effect or potential harm, not just biomarkers of exposure.

This just highlights some ongoing work of the group. The

3-HPMA inter-lab comparison study I've discussed. The second one on here, the proficiency test, it's the NNK metabolite NNAL that is being looked at.

One of the things the group is looking at is ranking of biomarkers for testing, in view of sort of their prioritization, and developing a guideline for reference standards.

CORESTA itself, and the biomarker subgroup in particular, adds value in a number of ways. It provides global interdisciplinary expertise in different sectors of the industry, to focus on advancing scientific knowledge and providing leadership and coordination of inter-lab studies to recommend analytical methods.

And I'll leave you with, if you're interested in more information, please contact Dr. G. L. Prasad. The address is shown here on the screen, or his e-mail address is prasadg@rjrt.com. It's not as visible on the web. P-r-a-s-a-d-g@r-j-r-t.c-o-m.

Thank you.

(Applause.)

DR. DRESLER: And our next speaker will be Kerry Lane from Palm Beach Autism Institute.

DR. LANE: Good morning. My name is Kerry Scott Lane, M.D. I'm a retired anesthesiologist. My talk today is entitled Pediatric p53 Tumor Mutations and FDA-CDC Tobacco Aflatoxin Data Demand Urgent Industry and FDA CTP Transparency.

I was the movement, that is, I initiated the $300 million civil suit against the tobacco industry in 1998 on the basis that the industry had deliberately concealed the presence of the ultimate carcinogen, heat-stable aflatoxin, on tobacco since 1965. Aflatoxin was then regulated by the FDA on all other agricultural commodities except tobacco.

A recent December 2015 *New England Journal* article, *New England Journal of Medicine* article from St. Jude's Children's Hospital has yielded insight into cancer mutations in over 1,000 pediatric patients; p53 tumor suppression mutations were found in 56%, and 8% RAS mutations, which are produced by the fungal and positive laboratory control carcinogen aflatoxin. Internal tobacco industry documents showed that aflatoxin was found on tobacco in 1965, hence the basis for the U.S. Justice Department lawsuit. I wrote in Medscape in 1999 for FDA to regulate aflatoxin levels on tobacco. FDA legislation was passed to regulate tobacco in 2009, and aflatoxin levels per brand were to be publicly released 3 years ago. FDA still has

not released its critical public health data.

The FDA-CDC in October 2015, in the *Journal of Agricultural Chemistry*, examined aflatoxin levels in smokeless tobacco. Tobacco snuff/snus had levels at 0.27 ppb, half of the FDA level for consumption in milk by humans at 0.5 ppb. Last week, major league baseball banned smokeless tobacco among all baseball league players.

Typically, for tobacco-induced cancers, clinical onset is 20-plus years until we see the disease. Oral tobacco users have cancers appearing after only several years of use. The product is not combusted at the consumer level. Most oral tobacco cancers have p50 mutations -- p53 mutations. Therefore, it seems clear that aflatoxin is the prime culprit or participant in this oral tobacco cancer epidemic.

Biomarkers for aflatoxin exposure include the aforementioned tobacco data which has not been released by the FDA. Additionally, urinary albumin, aflatoxin, glutathione adducts, and aflatoxin ELISA blood tests may be correlated with actual aflatoxin levels on smokeless products; p53 mutational spectra data can also be used as a biomarker of aflatoxin exposure.

Biomarkers of oxidative stress can also be indicative of

tobacco toxin exposure as well. Nasal and bronchial lavage, post-consumption of a modified risk tobacco product, could be used to compare aflatoxin exposure in an ultra-low or zero aflatoxin product.

Turning back to the FDA-CDC study of October 2015, it would appear initially that the aflatoxin levels on smokeless products are low. However, as low as they might seem, they must still be adequate to cause oral, esophageal, and other cancers, as evidenced by the clinical and epidemiological data.

At 0.27 ppb found in the FDA-CDC study, this means that at 0.27 ppb there may be thousands of molecules of aflatoxin in a wad of chewing tobacco, each one causing oxidative stress and disrupting mitochondrial oxidative phosphorylation and causing cancer. Apparently, this is enough to allow runaway tumor growth as the p53 tumor suppressor is silenced by aflatoxin. Nitrosamines may play a role here as inhibitors, but incapacitation of the p53 tumor suppressor gene by aflatoxin is a critical necessary step for the cancer to progress.

Technology exists in my patent to reduce the levels of aflatoxin on tobacco to essentially zero after 3 years, as is the case with the peanut industry. The FDA should mandate use of this aspect of my patented technology on all tobacco

products, as it is the cheapest and most efficacious method to essentially eliminate aflatoxin from all tobacco products.

I am offering my technology to the industry to produce a modified risk tobacco product -- and there are no applications to date -- in an expeditious manner and jumpstart the goal of the FDA CTP to produce a less harmful tobacco product.

If FDA is going to put a regulatory framework in place, it must be comprehensive, complete, timely, and hold the industry accountable for the role of aflatoxin in human carcinogenesis. FDA refuses to release its aflatoxin data, and that was precisely this type of concealment of tobacco toxicology that led to the United States Justice Department RICO case against tobacco. Over 150 RICO counts resulted, and one might ask, is history repeating itself?

In summary, the recent work at St. Jude's Children's Hospital indicates a majority of pediatric and 60% of adult tumors have p53 and RAS mutations. Aflatoxin and its likely combustion product, benzopyrene, are carcinogens known to mutate p53.

The FDA's 3-years-overdue failure to release this legislatively mandated data is perplexing. Children and adults are still being exposed to toxic tobacco products while the FDA

is silent with respect to the aflatoxin data and any methods the industry is using in an attempt to remedy this problem. I'm willing to collaborate with the industry and the FDA to solve this pressing issue.

Does anybody have any questions? I guess there's still a little bit of time left.

(No response.)

DR. LANE: Thank you.

(Applause.)

DR. DRESLER: Okay, so the two other speakers have not arrived or are not here. We will proceed to Session 4.

Thank you to the two public commenters for that morning opening session.

The next session, 4, is on cancer, and the first speaker will be Dr. Stephen Hecht from the Masonic Cancer Center at the University of Minnesota, and he will be speaking on Biomarkers of Potential Harm Associated with Cigarette Smoking.

DR. HECHT: Okay, I'm going to discuss biomarkers of potential harm associated with cigarette smoking, and this, of course, is related to cancer.

So this is an outline of my presentation. First, I'm going to give some background on co-carcinogenesis, tumor

promotion, and inflammation in tobacco carcinogenesis. Then I'll talk about some biomarkers of inflammation and oxidative damage, some biomarkers of exposure, which are also biomarkers of potential harm, and then genotypes modifying nicotine and carcinogen dose, and finally some comments about DNA adducts.

So this is the accepted central mechanism of tobacco carcinogenesis, in which carcinogens in cigarette smoke are metabolically activated by P450 enzymes and bind to DNA, forming DNA adducts which can cause miscoding. And when miscoding happens in critical regions of critical genes such as RAS and p53, the result can be loss of normal cellular growth control processes and ultimately the development of cancer. And this mechanism is supported by the thousands of mutations that are seen in the lungs of smokers at 10 times higher concentrations than are seen in nonsmokers.

But there are other things going on at the same time, and in this presentation I'd like to focus on co-carcinogenicity, tumor promotion, and inflammation.

So this is the initiation-promotion-progression model of mouse skin tumorigenesis, which was essentially bedrock in carcinogenesis research in the second half of the 20th century, and it involves initiation by a compound such as benzopyrene or

a nitrosamine or a nitrosourea forming DNA adducts. And then the genetically altered cell is promoted by other agents and ultimately goes through a progression stage leading to skin cancer, which was the main type of cancer investigated with this model.

So in this initiation-promotion protocol, which was originally developed by Berenblum and Shubik in the late 1940s, you paint mouse skin once with an initiator, typically benzopyrene or dimethylbenzanthracene, and then look 3 months later. If that dose of that initiator is low, you will see no tumors and no change. Or if you make multiple paintings with a tumor promoter such as tetradecanoylphorbol acetate, also you will see no change because these compounds are not carcinogenic by themselves. But if you paint once with the initiator and then give multiple paintings with a promoter, you will see tumors. So this is the classic initiation-promotion protocol, which was used mainly in mouse skin tumorigenesis but has also been used in other models.

And the classic tumor promoter is tetradecanoylphorbol acetate, which is a constituent of croton oil. And tetradecanoylphorbol acetate binds to protein kinase C and turns on all kinds of downstream effects, including the

NF-kappaB pathway of inflammation as well as the AP-1 pathway and MAP kinase.

So there have been mouse skin experiments which investigated tobacco carcinogenesis with respect to the initiation-promotion model. And the extensive fractionation experiments were carried out by Dietrich Hoffmann and his group, and this demonstrated that the tumor-initiating activity of cigarette smoke condensate is concentrated in subfractions containing polycyclic aromatic hydrocarbons, or PAH, such as benzopyrene. But when they tested a mixture of 17 PAHs in these subfractions at the concentrations in which they occurred or even twice that concentration, no tumors were produced. But the activity was restored when the 17 PAHs were added to the total cigarette smoke condensate. And other experiments ultimately showed that this was due to the presence of both tumor promoters and co-carcinogens in cigarette smoke condensate. And catechol was identified by the Van Duuren group as a co-carcinogen. And weakly acidic tumor promoters are also present.

In the 1970s, there were also inhalation experiments carried out, which demonstrated the importance of tumor promotion. The extensive studies of Dontenwill and colleagues

used Syrian golden hamsters that were treated intratracheally with dimethylbenzanthracene and then followed by inhalation of cigarette smoke. And these hamsters had significantly increased incidence of larynx carcinomas compared to the animals treated with DMBA alone or cigarette smoke alone.

And in more recent studies, the cigarette smoke exposure enhanced A/J mouse lung tumorigenesis, which had been induced by NNK via an NF-kappaB mediated pathway. So we know from all of these experiments and from other data that co-carcinogenesis, tumor promotion, and inflammation are closely associated.

Another recent experiment showing the role of inflammation in lung tumorigenesis in an animal model was carried out by the Kassie group at the University of Minnesota, and they treated A/J mice either with NNK or NNK followed by lipopolysaccharide, or LPS, which is an inflammatory agent.

So on the left you see a lung, a mouse lung that was not treated. In the middle is a mouse lung treated with NNK. And on the right is a mouse lung from the group treated with NNK followed by repetitive injections of LPS. And you see the histology on the top right. Below, you see the histology of adenocarcinoma induced by NNK plus LPS and also the

infiltration of macrophages in the inflamed lung.

So inflammation is implicated in tumor promotion, co-carcinogenesis, and tobacco carcinogenesis overall. And we know from other studies that anti-inflammatory agents inhibit lung carcinogenesis by tobacco carcinogens. And chronic inflammation is also implicated in human lung cancer etiology, and COPD is a risk factor for lung cancer, and both are characterized by chronic inflammation.

And in the next presentation, Dr. Shiels will present evidence on several markers of inflammation that were associated with lung cancer risk, mentioning one of these is C-reactive protein, which was elevated in just about every presentation yesterday.

So C-reactive protein is a nonspecific biomarker, but it definitely rises in response to inflammation. And inflammation, of course, is a nonspecific process.

Another biomarker of inflammation that has been fairly widely used is the prostaglandin E_2 metabolite, or PGE-M. So prostaglandin E_2 is a product of COX-2 catalyzed metabolism of arachidonic acid, and it's the most abundant prostaglandin found in various human malignancies, including colorectal, lung, breast, and head and neck. And PGE_2 promotes

inflammation and pain, and it has a role in tissue maintenance and regeneration.

So the best way to quantify systemic PGE_2 is with PGE-M, or prostaglandin E_2 metabolite. And this can readily be analyzed by liquid chromatography-tandem mass spectrometry.

This just shows an analysis from one sample from our laboratory, where you see a clean peak for PGE-M and the internal standard.

So PGE-M has been associated with both colorectal and gastric cancer, and its levels are increased in smokers, but that data are actually quite limited. So we actually need more data on the effect of smoking per se on PGE-M levels.

Oxidative damage and inflammation are closely related. They're like twins. Oxidative damage causes inflammation, and inflammation causes oxidative damage. I'm not sure which comes first, but a very good marker of oxidative damage is 8-epi-PGF2alpha, which also is readily quantified by LC-tandem MS. And this is a chromatogram and showing the structure of 8-epi or 8-iso-PGF2alpha.

This compound is consistently higher in smokers than nonsmokers in multiple studies. Yesterday we heard about the Total Exposure Study where 8-epi was 42% higher in smokers than

nonsmokers, but it's also affected by BMI, and it's related to total nicotine equivalents, also to age and gender -- females tend to be higher -- and smoking duration. And 8-epi decreases upon smoking cessation.

Another biomarker in oxidative damage is urinary 8-hydroxydeoxyguanosine, but this is not a very reliable biomarker. First of all, the methods can be prone to artifact formation, although that has been pretty much eliminated now. There's a large database of published studies of 8-oxo, or 8-hydroxydeoxyguanosine, in the literature.

And the increases that are seen in smokers are frequently small and highly variable. For example, smoking cessation for 6 months resulted in only a 23% decrease, while a separate longitudinal study of 8-hydroxydeoxyguanosine showed a coefficient variation of 48%. So this biomarker is quite variable and, in my opinion, not very reliable.

Now, let's talk about biomarkers of exposure and potential harm. So going back to the central mechanism, which has received a great deal of attention, the biomarkers of exposure are metabolites of carcinogens, and these can be modified to some extent by genetics. And there are also biomarkers of DNA adducts, which I'll touch upon. And less has been done on

reliable biomarkers of downstream molecular and biological changes.

So in the meeting in August on biomarkers of exposure, Jian-Min Yuan from the University of Pittsburgh presented data from the Shanghai Cohort Study that he leads, which showed that three biomarkers of exposure were also biomarkers of potential harm.

The Shanghai Cohort Study enrolled more than 18,000 men in the 1980s and carried out an in-person interview for smoking and other lifestyle factors, and they collected blood and urine samples at that time. Then they followed these individuals for incident lung cancer cases, which were identified through follow-up interviews and record linkage.

And ultimately, urine from 475 lung cancer cases and 475 matched controls -- they're all smokers -- were analyzed in our laboratory, and we evaluated total NNAL, phenanthrene tetraol or PheT, total cotinine, and mercapturic acids with respect to lung cancer in the study. And the results showed that total NNAL, phenanthrene tetraol, and total cotinine were all independently related to lung cancer, even after adjustment for smoking duration and intensity and the other biomarkers.

This was not the case with the mercapturic acid. When you

adjust for total nicotine equivalents, the effect of the mercapturic acids disappear. But total NNAL, phenanthrene tetraol, and total cotinine were all biomarkers of potential harm in the Shanghai Cohort Study. So they are biomarkers of exposure, but they are also biomarkers of potential harm.

Now, total nicotine equivalents and NNAL can be affected by genetics, and one of the main variants that has been studied is the genetic variation in the alpha5 nicotinic cholinergic receptor subunit, or CHRNA5, on chromosome 15q25, which is associated with the heaviness of smoking and nicotine dependence.

And there are two highly correlated genetic variants shown here, and these variants lead to higher nicotine and NNK uptake by changing the conductance in the nicotine receptor. And the A allele, the rs16969968, was associated with a lower likelihood of smoking cessation and a 4-year earlier age of lung cancer diagnosis.

This has been diagrammatically represented by the Beirut group from WashU. Basically, the AA genotype leads to a 4-year delay in median age of quitting and a 4-year earlier median age of diagnosis of lung cancer.

Another modifier of nicotine metabolism is the CYP2A6

genotype. CYP2A6 codes for cytochrome P4502A6, which is the major enzyme involved in nicotine metabolism. And many variants -- I think it's about 80 -- of CYP2A6 have been identified, and these variants affect nicotine metabolism, nicotine dose, and lung carcinogenicity.

You see a reduced rate of nicotine metabolism associated with lower cigarette consumption and lower dependence scores, lower brain response to smoking cues, and greater cessation. All of this is because you don't have to take up so much nicotine because you have more on board when you have a deficient metabolism with nicotine.

My colleagues in our program project grant have looked at the CYP2A6 genotypes in five different ethnic groups with differing risks for lung cancer. I don't have time to go into all of the details, but if you just look at the Japanese Americans on the right, the red and light-red pie segments indicate the low activity forms of CYP2A6, and you can see that those are highly more prevalent in Japanese Americans.

And looking at the patterns of nicotine metabolism in these five groups, which this comprises about 85% of all the nicotine metabolism, including cotinine, 3-hydroxycotinine, and N-oxides, the glucuronides, and the free nicotine as well as

nicotine N-oxide, you will see that in the Japanese Americans, the red pie, which is free nicotine, is greater than in most of the other groups.

So because the Japanese Americans have more deficiency in CYP2A6 activity, they have more nicotine on board, therefore need to take in less nicotine and therefore smoke differently. So they're taking in less nicotine and less carcinogens, less tumor promoters, and less of everything else, and therefore they have a low risk for lung cancer.

The CYP2A6 genotype and nicotine metabolism has also been investigated in the Shanghai Cohort Study. And on the left are lists of various different CYP2A6 genotypes, going from normal metabolism down to low metabolism.

And the subjects with each of these genotypes were phenotyped using the ratio of 3-hydroxycotinine to cotinine, which is a phenotypic measure of CYP2A6 activity. And you can see that the normal metabolizers had the highest ratio. The intermediate metabolizers had a lower ratio than the slow metabolizers and the poor metabolizers. So this ratio, it's a very good indicator of CYP2A6 genotype and phenotypic CYP2A6 activity. They go together very nicely. And this was looked at with respect to lung cancer in the Shanghai Cohort Study,

and the result was that when you separate normal, intermediate, slow, and poor, there was not a significant trend for CYP2A6 genotypes and lung cancer.

But if you combine normal, intermediate, and slow and compare that to poor, there was a decreased -- a significantly decreased risk for lung cancer in the poor metabolizers compared to the normal, intermediate, and slow. It wasn't a tremendously strong effect, and when you correct for total nicotine equivalents, it actually loses its significance, but these were the results. So CYP2A6 genotype modifies nicotine metabolism and may itself, to some extent, be related to lung cancer incidence.

Now, let's return to DNA adducts and the central mechanism. And I've already mentioned the biomarkers that are metabolites. Now looking at the DNA adducts, and the DNA adducts are very important because DNA adducts lead to mutations. So a person may have high exposure but low metabolic activation. That person is going to have, relatively speaking, lower DNA adducts than one might have expected. Or you might have relatively low exposure but high metabolic activation, so you might have higher DNA adducts than one might have expected. So the combination of DNA adducts and

biomarkers of exposure could be very important.

But DNA adducts are very challenging to measure. They're present in extremely low concentrations. And starting basically in the 1980s through the early 2000s -- and there are still a few studies going on -- DNA adducts have been quantified by ^{32}P-postlabeling and immunoassay. There are many, many studies, but these studies are non-quantitative because the methods don't give you quantitative data on the DNA adduct levels. And the adduct structures are basically unknown in most of these studies. So these are big weaknesses when you try to interpret the data.

Many studies have shown higher adduct levels in general in various tissues of smokers versus nonsmokers. For example, the smoker's lung had the so-called diagonal radioactive zone that's seen in postlabeling studies. And this is higher in smokers than nonsmokers, but nobody knows what the adducts are. And there have been some meta-analyses trying to associate DNA adducts, measured in this way, with lung or bladder cancer risk, but they didn't show much. We have used oral cell DNA and looked at tobacco-specific HPB-releasing DNA adducts by LC-tandem mass spec, and we see relatively high concentrations of these HPB-releasing DNA adducts in the oral cells from

smokers. These adduct levels are actually similar to what one sees in animals treated with carcinogens. So this requires further study. It's a good lead.

And mass spec methods are becoming more and more sensitive. If you look at this slide -- and this is from an animal study -- in the top two panels, A and B, we were barely able to detect O^6-pyridyloxobutyl-deoxyguanosine in esophageal DNA from rats treated with NNN.

But in the bottom two panels, C and D, we could easily detect and quantify these using liquid chromatography, nanospray ionization, high-resolution tandem mass spectrometry.

So the nanospray ionization concentrates the sample, increases the sensitivity, and high-resolution mass spectrometry gives you a great deal of selectivity to get very clean, sensitive, quantifiable data. And we've been able to use this method to detect formaldehyde DNA adducts in human oral cell DNA.

And a number of other groups have actually looked at oral cell or salivary DNA for a variety of different DNA adducts. And so this part of the science is really evolving rapidly and, I think, has a great deal of promise.

So, in summary, inflammation, co-carcinogenesis, tumor

promotion, and oxidative damage are important in tobacco carcinogenesis.

Some biomarkers of inflammation have been associated with lung cancer risk.

8-epi-PGF2alpha is a reliable biomarker of oxidative damage in smokers, but its relationship to harm is unclear right now.

Total NNAL, phenanthrene tetraol, and total cotinine are related to lung cancer risk in the Shanghai Cohort Study.

The CHRNA5 and CYP2A6 variants modify nicotine and carcinogen uptake in smokers and are related to lung cancer risk.

And DNA adducts quantified by mass spectrometry are possible new biomarkers of potential harm.

Thank you.

(Applause.)

DR. DRESLER: Thank you, Dr. Hecht.

Our next speaker is Dr. Meredith Shiels from the NIH National Cancer Institute, and she will be speaking on Cigarette Smoking and Variation in Circulating Markers of Inflammation and Immunity.

DR. SHIELS: Good morning, everybody. And thank you to

the organizers for the opportunity to come today and share some of the work that I have done in the Division of Cancer Epidemiology and Genetics at the National Cancer Institute.

As it is no surprise to this group, cigarette smoking causes cancer at about 18 different sites, as well as a range of other chronic diseases, including coronary heart disease, stroke, and COPD.

Smoking-induced inflammation and immune modulation are emerging as potential important mechanisms in the development of cancer and other chronic diseases.

And we know that smoking leads to a number of pulmonary and immunological changes, including studies that have shown that smoking alters the prevalence and function of many immune cell types in the lung, as well as altering circulating levels of a number of inflammation markers, which is what I'm going to concentrate on today. There have been a couple of studies in the literature that have examined circulating levels of immune markers among smokers. For example, the Total Exposure Study showed that smokers had increased white blood cells, C-reactive protein, and fibrinogen levels compared to nonsmokers. And these associations were held with cigarettes per day, smoking duration, and nicotine equivalents in adjusted models.

Another study, recently published in the VITAL cohort, examined a handful of inflammation markers in circulation, including CRP again, prostaglandin, soluble tumor necrosis receptor 1 and 2, TNF-alpha, IL-1B, IL-6, and IL-8.

Just to highlight the significant results here, this study again found that those who smoked more than -- had a smoking history of more than 18 pack-years had significantly elevated levels of CRP compared to never smokers, and that prostaglandin levels were also elevated in this group. And again, these associations held in adjusted models.

Now, though there have been a few prior studies examining markers of inflammation and circulation in relation to tobacco use, most of these studies have been limited to a few markers, as I just showed you, primarily CRP and IL-6, but also a few others.

I had the opportunity to examine a broad panel of the inflammation, immune, and metabolic markers in the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, using data from three already conducted case-control studies of cancer and their controls.

For those of you who aren't familiar with the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, or PLCO,

it was a cancer screening trial that recruited approximately 155,000 men and women age 55 to 74 years old during 1992 through 2001. The PLCO collects demographic, behavioral, and dietary information, and in the screening arm, it also collected blood samples at baseline and at five subsequent annual visits. Cancers in PLCO were ascertained through annual questionnaires and confirmed by medical chart abstraction and death certificate review.

I thought it might be of interest to understand how smoking information was collected in PLCO, for this audience. So you can see that smoking status was ascertained through asking participants about if they had regularly smoked for 6 months or longer and whether they currently are regular smokers. Duration was assessed by asking of the age at which they started smoking. Intensity was assessed by asking, during periods when you smoked, how many cigarettes did or do you usually smoke per day? And finally, time since quitting was ascertained by asking, at what age did you last stop smoking cigarettes regularly?

I want to go over the study design for this particular analysis because it was a bit complicated. So as I mentioned, the study is nested within the PLCO screening arm, which is the

arm that collected blood samples.

There were three initial studies aimed at examining a wide panel of inflammation and immune markers in relation to cancer that were nested within PLCO. They included a study of lung cancer, which I led, that had 526 cases and 592 controls, a study of non-Hodgkin lymphoma with 301 cases and controls, and a study of ovarian cancer of 150 cases and 149 controls.

The data from these three studies were combined, and we used these data to examine a variety of exposures in relation to inflammation marker levels. Because these studies matched on different criteria and included participants with different characteristics, it was important to then reweight the data from each of these studies to the full PLCO cohort so that our results are representative of what would be seen in the full PLCO cohort had we measured the inflammation markers on all of the participants. To do that, we had to exclude non-whites, those who had cancer prior to randomization, and those with missing smoking information.

So we then reweighted our dataset of 1,819 people to the full 58,264 people in the PLCO screening arm that met our inclusion criteria.

Each of these case-control studies measured inflammation

markers using a Luminex bead-based assay that measures multiple markers simultaneously. This includes markers with many functions, including pro- and anti-inflammatory cytokines, chemokines, soluble receptors, and metabolic markers. And it requires a relatively small amount of specimen (420 µL) to measure a large number of markers. And there was a prior methods study carried out in my division to assess the performance and reproducibility of these markers.

This slide just shows the 86 markers measured in each of the individual case-control studies that contributed to the analysis that I'm about to present.

So, first, to show you the characteristics of the participants in the study, the first column here presents the characteristics of the individuals that had inflammation markers measured. The second column presents the weighted number of people and characteristics after statistical weights were applied. And the third column presents the characteristics of those in the PLCO screening arm.

You can see that though the characteristics of those measured in our study weren't exactly representative of the PLCO screening arm, after applying weights, they were. So for example, and maybe most pertinent is if you look down here at

the smoking status in our measured -- in our group of measured values, 23% were current smokers, and that's largely because we included a study of lung cancer that had mostly smokers involved. But after weighting, current smokers only represented about 10% of the data, which is equivalent to what's seen in the full PLCO trial.

This slide presents the main results of our study. So here we're comparing current to never smokers, and the outcome of our study is above versus below the median level of each marker. We used weighted logistic regression and adjusted for a number of factors. Of the 86 markers, we found 19 markers with a p-value of 0.05, and after correcting for multiple comparisons, 10 markers passed false discovery rate criterion. And those are shown here. So this figure is showing odds ratios, again comparing current to never smokers, with the outcome being above versus below the median value of the marker.

You can see that, of our 10 markers, the majority were actually reduced in smokers, implying an immune-suppressive effective of current smoking, and that included two soluble receptors, the VEGF receptor 3, and the IL-6 receptor, stem cell factor, and four interleukins, so interleukin-16, IL-1B,

IL-1RA, and IL-15.

In contrast, two chemokines were elevated in smokers: CCL11 or eotaxin, and CCL17 or TARC. And as has been mentioned many times previously, C-reactive protein levels were also elevated in smokers compared to never smokers.

We then examined former smokers and compared the inflammation marker levels in former to never smokers. And the red dots represent the odds ratios, and the lines represents the 95% confidence intervals in the former versus never smoker comparison. You can see, for all of the markers, the associations are attenuated, with only four remaining statistically significant and that's for soluble IL-6 receptor, stem cell factor, IL-16, and IL-1RA.

Among current smokers, we then examined intensity or cigarettes per day and duration or number of years smoked. We found no associations with cigarettes smoked per day among current smokers.

Now, it's important to note that cigarettes smoked per day was assessed in PLCO in categories. So the lowest category was 0 to 10 cigarettes, so we couldn't ascertain whether or not there is some dose response within that lowest category. That was the lowest category that we could examine. And we only saw

associations with smoking duration for stem cell factor and soluble IL-6 receptor.

We then looked at former smokers, and we found that for six markers, there were significant trends across categories of time since quitting, including CCL17, CCL11, IL-15, IL-1B, IL-1RA, and CRP. And for these markers, the levels appeared to be similar to those of never smokers within 5 to 20 years of quitting. And I'm going to show you a little bit more information on that with these figures.

So here, what I'm showing you is odds ratios again of having high versus low levels of the markers across categories of time since quitting among former smokers, with current smokers as the referent group.

So this is for CCL17. You can see a significant trend with years since quitting, starting with current smokers as the reference. In the highest categories of years since quitting, the marker levels actually look similar to what you see in never smokers.

And I should note that I'm not going to show every figure for every marker, but similar patterns were observed for CCL11 that looked very similar to this figure here.

With a marker going in the opposite direction, IL-1RA

marker levels, the odds of having higher marker levels increased with increasing years since quitting, again compared to current smokers with a significant trend, and again approaching similar levels to what are seen in never smokers with an increased number of years since cessation. And we saw similar patterns for IL-15 and IL-1B, as what's shown here.

CRP was a marker that had sort of a strange trend in that in recent quitters the levels of CRP were actually significantly higher than in current smokers, before declining significantly over time. And we think this is likely confounding due to indications. So those who may be sick may have cardiovascular disease, and therefore higher levels of CRP are more likely to be recent quitters. They're quitting because of an underlying condition that may be increasing their CRP levels.

For four markers, we saw no significant trend with years since quitting for soluble VEGF receptor 3, IL-16, soluble IL-6 receptor, and stem cell factor. And these markers appeared to remain at an intermediate level between current and never smokers, even after many years since cessation.

So I'd like to just point out a few of the limitations of the study and also opportunities for future research that may

build on what we found. So first, this is a population of older, longer-term smokers. The population at enrollment was between 55 and 74. This doesn't really give us -- may not give us direct insight into what's happening with younger smokers.

Smoking characteristics were limited both by the questionnaire and also by the people who were included in the study. So, for example, we saw no association between smoking duration and marker levels, but the duration range was very limited in this population. The age range was very narrow at enrollment, and all of the smokers appeared to have started smoking around the same age, so we didn't have that much variation in our number of years smoked.

As I mentioned when I presented the results for intensity, the number of cigarettes per day was collected categorically, so we weren't really able to get at any sort of dose response that might be seen between 0 and 10 cigarettes per day because that was the lowest category.

This is a cross-sectional design. And so when I'm presenting associations about, for example, years since quitting, I'm comparing those who quit 5, 10, 20 years ago to a different set of people who haven't quit or who are never smokers. So these aren't the same people being followed over

time, so there's a need for studies that look at serial measurements over time among the same individuals.

And with all of our studies that measured circulating markers of immunity and inflammation, we need to remember that this is measured in circulation. And if you're interested in local inflammation in the lung, it's not clear whether the levels in circulation are representing what we see in the lungs.

And finally, this study needs to be replicated. This is the only study that has measured such a large number of markers related to smoking using this panel. And while the technology is great for allowing us to measure a lot of markers in a small amount of specimen, these associations do need to be replicated in other studies.

So as Dr. Hecht mentioned, I'd like to now just talk a little bit about cancer, since I am from the NCI.

CRP, which is one of the markers we identified here, has been consistently associated with an increased lung cancer risk, both in our work and other studies as well as other cancer sites.

But it's important to remember that a lot of the studies that have examined these markers in relation to cancer risk

have really tried to eliminate the effect of smoking on the marker-disease association. So, for example, the study that Dr. Hecht cited of mine and the original study that that study replicated measured all of these markers and found associations with lung cancer risk.

But the goal is not to understand the mediating effect of these markers in relation to smoking. The goal of the study was really to control as closely as we could for smoking and then assess the association between inflammation and lung cancer risk. So we really need a different study design to get at the mediating effect of these inflammation markers between smoking and cancer risk.

Now, within our datasets and PLCO, though these are closely matched lung cancer studies, we do have the opportunity to statistically break the matching and reweight the studies to the full population and try and get at this mediation question, which it's really important to understand whether these inflammation markers are mediators of the smoking effect on cancer risk. And that's something that is ongoing at NCI.

So, in summary, we observed associations between smoking and 10 markers of inflammation and immunity with various functions, including markers associated with chemotaxis,

inflammation, anti-inflammation, cell development and differentiation, cell growth and activation, angiogenesis and hematopiesis. These results suggest that the effects of cigarette smoking are largely immune-suppressive, which is consistent with some reports that nicotine is immune suppressing.

There's a low threshold for the effect of smoking on circulating immune markers, and that we did not see a dose-response relationship, but it implies that even a very small exposure to cigarette smoke, 10 cigarettes or less, has an effect on immune markers, though as I mentioned, there is the caveat that we weren't able to look at lower numbers of cigarettes per day in relation to a dose response with marker levels.

Levels of C-reactive protein, chemokines, and most cytokines approached levels in never smokers with increasing time since smoking cessation, implying that markers may gradually revert back to normal once exposure to cigarette smoking is removed.

In contrast, levels of soluble VEGF receptor 3, IL-16, stem cell factor, and soluble IL-6 receptor were not related to time since quitting. And these markers may remain at

intermediate levels between never and current smokers for many years.

And finally, more research is needed to replicate these results longitudinally in these same individuals and to understand how inflammation markers may mediate the association between smoking and cancer.

And with that, I'd like to acknowledge my collaborators at the National Cancer Institute, Leidos Biomedical Research, which is the laboratory that carries out all of our inflammation marker work, FDA, the University of Maryland, and our contractor, Information Management Services. And finally, I'd like to thank the funder, which is the Intramural Research Program at the National Cancer Institute.

Thank you.

(Applause.)

DR. DRESLER: Thank you, Dr. Shiels.

Our next speaker for this morning's session will be Epigenetic Biomarkers of Tobacco Smoke Exposure by Dr. Douglas Bell from the NIH National Institute of Environmental Health Sciences.

DR. BELL: Okay. Well, good morning. Thank you for this invitation to come and talk about our work. I am funded

primarily by the Intramural Program of the NIEHS, and we have some funding from the FDA Center for Tobacco Research.

So today I'd like to talk to you about a new area, a new type of biomarker, a potential biomarker, and share with you some of our preliminary research.

So to start off, we're talking about epigenetics, and I suspect not many of you are very familiar with epigenetics. So we're talking about something that's happening -- let me see if I can get this thing to work. Yes. So we're talking about things that are going on in the cell nucleus, in chromosomes, and then at the level of DNA. So DNA can have these methyl groups, and I'm going to be talking about DNA methylation for this study.

So if I can get that to go forward. There we go. Okay. So as we've just heard from Steve Hecht, we know very much that smoking can cause damage to DNA, can cause mutations, and it can cause toxicity, and this is a very established pathway for many different negative outcomes from smoking. But of course there are many things that are not understood about the exposure to disease process.

And so one of the -- this area that we're looking at is, do epigenetic changes tell us something or is it a new

mechanism for how these adverse outcomes occur? And the key sort of idea is that these epigenetic changes, which could occur during your lifetime, during your exposure lifetime, may turn genes on or off, and these on/off events lead to some adverse outcomes such as co-carcinogenesis, which was talked about earlier, or tumor promotion. Those things could be caused by genes being turned on via this type of mechanism.

So again, the DNA can be methylated, and it typically occurs at these sequences called CpGs. There's about 20 million of these in the genome. So they can be converted to hydroxymethyl C's by these TET enzymes, and then they can be essentially demethylated through another mechanism.

So the point simply is to show that they can be methylated and demethylated, and this happens in cells as they -- sometimes as they divide, sometimes as they differentiate, and sometimes as a result of exposures, in other words, a signal.

So to jump down to the DNA level in a real example, we're talking about a gene called the AHRR gene, aryl-hydrocarbon receptor repressor gene, and this just shows a cartoon of its structure. And then if we look at a genome browser view -- so we're taking a really close-up view and this -- let me see. This shows that this area is within what we think is a control

area for this gene. And we can use arrays.

So I'm going to just mention several methods for measuring these CpGs, for measuring the methylation. So we can use arrays, which will pick out some specific regions across this gene or across the genome, or we can use total whole genome sequencing.

So in this case, these bars represent individual cytosines, and these are all methylated or -- yes, fully methylated. And then these ones down here, which show the low level, are unmethylated. So this is an example of displaying all of the hundred or so of these cytosines in this region of the genome.

And then we've used different kinds of methods to identify regions that are variable. We call them -- sometimes we call them DMRs, differentially methylated regions. And we can use -- so this is a technique called reduced representation by sulfate sequencing, which doesn't get every CpG in the region; it only gets a subset, a reduced representation.

Notice I've put the prices on here. These arrays are around 300-and-some dollars. Whole genome sequencing ranges around $3,000 to $10,000 these days, and this technique is about $350. And then there are various PCR-based methods where

you can just look at a small area of the genome, and they're much less expensive.

So we've been using all of these techniques in trying to verify that the results from one method agree with the results from another method. And I'll present just a little bit of that data.

So to get back to why this methylation is important, I made some diagrams for you. So this shows a gene that's turned on. It's a small little diagram of the gene, and it has a polymerase there that's getting ready to make RNA and transcribe it. Well, years ago everyone believed that the methylation level -- and this is supposed to be a depiction of methylation in several different samples. The methylation level of the start site was very important, the promoter was very important.

Now, it turns out we've discovered, in just a few recent years, that it turns out there are these enhancer regions which are really the control mechanisms, and the methylation of the enhancers is really the key. So we see that the -- and this shows these color codes with the samples, that if you have low methylation here, it's more likely related to high gene expression.

And I have another picture here which sort of displays it in this stick diagram method. So it turns out these enhancers are turned on when they have a lot proteins binding here, and the event of them binding is associated with the change in methylation level. So we're trying to find these associated with smoking.

One of the other features of this epigenetics is that DNA is either wound up in closed formation or open for expression, and this diagram suggests that, is that these control regions, when they're bound with lots of DNA-binding proteins and polymerases, are in the open and on position. When they're wound up tight and methylated, they're in the closed position. So this is this on/off aspect of what we're looking at.

All right. So to go to actual data from a population, we've been measuring this methylation in blood samples of smokers and nonsmokers. And I'm going to show you a graph of what's called an EWAS study, an epigenome-wide association study.

So this shows significance on the y-axis here, and then it -- there were 450,000 CpGs measured for methylation in this experiment, and they're all plotted here, and then this basically shows the p-value for the association with smoking.

So we see that, and we use this stringent statistical cutoff of p 10^{-7} or so, which tells us what the -- what we call genome-wide significance, and that's what the red arrow indicates here. This is 10^{-7}.

So all of these, about 700-something, are very significant, associated with smoking. And I've circled some of the ones that have been identified, and I think it's more than 15 studies now have identified that these particular cytosine methylation sites are altered in smokers versus nonsmokers.

This is actually a very remarkable thing, that a technology that just sort of came out of the box is reproducible by pretty much everyone that's got a sample of smokers and nonsmoker DNA. And, in fact, these samples were in my freezer for 20 years, and we see the same result in these samples as all of the newly collected samples. So it's quite remarkable. And DNA methylation is stable over time in storage anyway.

So I'll be talking a little bit about this AHRR gene, and in that gene there's one highly significant marker, a CpG marker, and then there's a whole other series of them in the gene.

Okay. Well, one point that's important is that these

samples were measured with whole blood DNA, so that's a mixture of cells. And I'll come back to that concept.

What about specificity in specific cell types? So we have some questions here, and I hope I'll answer some of them. We do know that these changes are reproducible, at least qualitatively reproducible. The quantitative studies have not really been done. So are they really dose dependent? Is it a dose-dependent biomarker of exposure? How does this compare with other biomarkers? So I'll talk a little bit about that.

So these two graphs plot this methylation level of this specific one, the 5921, versus smoking dose. So this kind of shows you've got a strong -- there's a highly significant relationship with dose because all the nonsmokers are up here and the smokers are sort of spread out in this cloud. And if you look at pack-years, kind of a similar thing, they're spread out a little bit better.

You can see that, as sort of the specificity thing, this group is pretty good. There's not too much overlap with the smokers and nonsmokers. However, actually, this is a group of fairly heavy smokers, but there's certainly plenty of scatter in this kind of measurement.

And this just shows it in another way with basically bar

graphs here, and we can see that the dose response is not really very dramatic when you get to the heavy smokers. I mean, there's a very -- there's a pretty good dose response for light smoking, but you sort of don't get any more effect with more smoking. And this seems at least partly related to the fact that we've got sort of this limited dynamic range, and we're measuring cells in blood, and it's a mixture of cells, some of which may be sensitive and some which are not. And I'll get back to that a minute.

So I want to compare this a little bit with some other markers. We've heard lots and lots about cotinine in the blood. Well, so we had a small, a very small sample in which we could measure several things in the blood and also measure methylation.

So this is just about 25, 26 people here. This shows their average cigarettes per day, years of smoking, and the cotinine level. So as you know very well now, cotinine is quite a nice marker of current smoking. These nonsmokers have really no cotinine.

Now, this is another biomarker that Steve Hecht didn't really discuss. It's a protein biomarker, and it's quite a nice marker of smoking exposure, and you can see that these

levels are quite high in the smokers. So we had this small study, and we wanted to compare this with our methylation. So let's see. This displays the relationship between one of these protein markers, here, the HEVal adducts, and this is the cotinine levels.

So you can see that there's -- although it's kind of an odd distribution of individuals, these are one-pack-a-day smokers here. You can see that there is an increase in adducts, of these adducts in the smokers. And it just turns out that this particular adduct is metabolized through a polymorphic glutathione transferase, and the individuals that are null for this transferase have sort of higher levels of adducts versus those that have this.

However, this is just the -- this was published quite a while ago, and I wanted to show you then the relationship between these adducts and methylation.

So again, in a pretty small -- well, this doesn't show the numbers, but it's just 25 samples. You can see that the relationship between these two adducts and smoking is on the order of -- the R^2 is about 0.58, 0.6, so explaining about 60% of variance. Pretty strong p-values here. Cotinine is similar.

So for our CpGs, here's two of the strongest or the most responsive CpGs. We see a similar significance level. I'll come back to some of the other issues with this.

So this is the levels in whole blood, measured in whole blood. Okay, so moving on. So again, with whole blood. Really to understand this a little bit better -- and again, we don't really understand the mechanism of what's going on here at this point.

Are the changes similar in all different cell types? And are some cell types more sensitive? How about the methylation change? As I mentioned, the mechanism should be that there's a gene expression change. So can we actually demonstrate that? And how are these differentially methylated regions? How can we identify whether these are actual regulatory regions?

So this shows some of the different cell types in blood, all right? And hematopoiesis proceeds from these stem cells to progenitor cells to all of these various differentiated cells. And there are many, many, many more possible types of differentiated cells. So we're kind of asking, you know, is the effect of smoking on sort of all the cells at the same time? Is it on the progenitor cell and then that's sort of transmitted to the rest of the cells? And so we're doing a lot

of separations on these cell types. I'm going to compare these two cell types, T cells and monocytes.

And just to be clear, even within the monocytes there could be different things going on. There could be some other cell types. There could be just a group of cells that are changing. So, in fact, we can know that that actually must be the case because an individual cell is either methylated or not. And then we measure some kind of a 50% change, so there must be cells in there that aren't changing. And anyways, that's an important feature. So I've shown just some of the cells have been sort of altered. And then if 3 out of 10 cells are altered, then you get a 30% methylation level.

So let's see. So which cells are most sensitive? I'm going to go compare these T cells and CD14 monocytes and also look at a small group of individuals that we collected at NIEHS.

So this is from the Multi-Ethnic Study of Atherosclerosis, and I was happy to meet one of my collaborators on this study today or yesterday -- last night at dinner. So that's one of the nice features about these meetings. So I'll talk about the result from the MESA study.

And so when we look at 114 current smokers versus never

smokers and we -- this is one of the EWAS or this is a regional epigenome association plot in which we -- just looking at the region around the AH receptor repressor gene, and each one of these circles represents the significance of a specific CpG. And so we see that there are some very significant associations with smoking, and they show about a 25% reduction in methylation of this gene here. Highly significant, 10^{-134}.

And when we look at T cells, CD4 T cells, helper T cells, we also see highly significant association between some of the CpGs in this gene. But the association is less and the difference is less, as well. So the sensitivity of these T cells is less to this exposure.

So this is quite interesting. This shows a relationship with former smokers and current smokers. So we looked in a large group of former smokers, and this shows the level of methylation. The level changes are relatively small, but the significance is very high, so that even many years later you can detect a significantly changed, altered methylation in these former smokers. And you can see, though, that the change is quite a bit greater in the current smokers.

So this gives you a plot of the former smokers and with years since quitting here, and that's kind of interesting. The

current smokers are down here around 0.5-something, and then fairly quickly after you've quit smoking, it's going up pretty much. But it sort of hangs on the methylation difference from nonsmokers, sort of hangs on for many years.

And one of the interesting things about this biomarker, though, is that there do seem to be some outliers here and that there are some people that have this altered methylation even -- well, I would say these people look like current smokers. And so the question is whether or not those people still are smoking, and they just say they're quitting, they've quit, or whether or not there's some other aspect that is driving the methylation down in these individuals.

So in our study from NIEHS, just a small group, we split out multiple cell types. First of all, it shows the result in whole blood for this one CpG, and you get about a 20% difference, and in isolated peripheral blood mononuclear cells you get a little bit less change. And in monocytes and granulocytes, you get a pretty big change, 25, 30%, and less of a change in B and T cells. So that's a very interesting result, and we're trying to work on why this is the case.

I'll show you a little -- a result of sort of a display of three different markers. And this is kind of interesting.

This is the mononuclear cell level. Here's the monocytes and the granulocytes, and you see that the change -- this just plots the change in methylation. The changes are sort of similar in the myeloid cells and then lower in the lymphocyte-type cell.

So the patterns kind of fit each other. So we're thinking that this is sort of a feature of the biology here. So the myeloid cell methylation levels are highly correlated. And I'll remind you that the numbers of cells in these groups quite vary the percentages. So granulocytes dominate the blood, and then we have T and monocytes and B cells. So changes in these proportions could have effects in the outcome of the measurements.

And I'll just show this rather quickly, that when we compare cotinine levels with the methyl levels in whole blood, for this small sample we get sort of modest significance, but in monocytes and granulocytes we get a much better association with smoking.

Now, to come back to this idea that we're measuring something, a group of CpGs in this gene, and we're looking at the stem cells here, and these red bars display the difference associated with smoking. So this area here is responsive to

smoking and -- all right. So when we look at multiple cell -- different cell types here, we want to see if the same kinds of changes occur in these other cell types. And this displays this browser image and shows the result for the stem cells, monocytes, granulocytes, and NK cells.

So this is using a sequencing technique called RRBS, or reduced representation bisulfite sequencing. And we can see that, in this region, all of these cell types show this sort of qualitative change in methylation associated with smoking, right? And so we've used some different techniques. I'm not going to go into this in very much detail, but we can quantify at these different locations and see that there's a difference between nonsmokers and smokers that's highly significant.

We use another technique called amplicon sequencing, and we also see that there's a significant difference for these. And we found some -- the arrays had this 5921, and with sequencing we found some additional things that were more differentially methylated than the array locations.

So this plots it in another way, showing the data for this 5921, smokers/nonsmokers, and this novel CpG, which the distribution isn't really very different, but the p-values are a little bit stronger. But there is overlap with some smokers

and nonsmokers appearing kind of similar.

This shows it related to cotinine. And since we're plotting cotinine on the X here, we see that the cotinine still very much separates smokers from nonsmokers here, but we see a good relationship with the change in methylation as well.

So we see association with gene expression, and this is with several -- 373 individuals, a highly significant change. And we see also the same kind of highly significant change in gene expression in our smaller sample.

I'm going to just finish up with the idea of whether or not this is a marker of exposure or of an intermediate marker of pathology or effect or whether it's a marker of susceptibility.

So in our study with MESA, we compared carotid plaque scores with AHRR methylation, and we see that there was quite a significant association between the methylation level of this 5921 and plaque scores.

And this shows the effect size. The significance is greatest with the combined current and former smokers. The effect size is higher in the current smokers.

Now, at this point we can't really distinguish, though, whether or not the methylation level is really just a better

measure of smoking than asking people if they're a smoker or not. So we can't really distinguish if we can put methylation in between exposure and the early signs of atherosclerosis.

So this slide just gives some of the summary of what we have here. I should say that, in general -- I'll just flash these up since my light is red. I'll just say that, you know, these studies are not really designed to validate biomarkers. So you might consider a different design, if that was the point. But it's really a measure of exploring some new biomarkers.

And we have some studies ongoing to try to do further association with risk. And so we're doing much more separation of different blood cell types, and we're working with the same PLCO study with Nat Rothman and Lee Moore to look at association with smoking and disease. And we're looking in a mouse model to look at if we can identify the same kinds of methylation changes.

So thank you very much for your attention. I'm sorry I went over a few minutes, but thank you.

(Applause.)

DR. DRESLER: Well, if you are looking at the agenda, we are scheduled for a break, but -- and I'm wondering. I know

the coffee kiosk is open, but people usually don't mind if we finish a bit early in the day.

So, Dr. Miller, would you be willing to go next? And then we'll move up the panel discussion and maybe we'll finish earlier today, if that's all right.

So our next speaker is Dr. York Miller from the University of Colorado School of Medicine, and he will be speaking on Airway Epithelial Biomarkers of Tobacco Risk and Exposure.

DR. MILLER: Okay, great. And the timer will go to 25 and I should stop, is that basically --

DR. DRESLER: You do, but you know, yeah --

DR. MILLER: Okay.

DR. DRESLER: -- you have to deal with me.

DR. MILLER: Okay. Well, thank you. I've found this to be a very interesting meeting to attend, and so I've learned a lot.

I'm a pulmonary doctor, and my interest has been really biomarkers of risk and chemoprevention as applied to lung cancer. I'm probably -- okay. And I'm going to start by talking about the -- what is thought of as the gold standard biomarker that's been used in a lot of chemoprevention -- in some chemoprevention trials, and I'm going to go through

exactly how good a gold standard that is. And at the end I'll just bring in some summary of chemoprevention results that have come out.

So I think the gold standard of saying that dysplasia is a biomarker of risk has largely been shaped by anecdotes, and if you look at the literature critically, it's really all over the place.

But here's an anecdote. Here's a patient that was treated for squamous cell carcinoma of the larynx. In 1982 he stopped smoking. He entered our SPORE bronchoscopy cohort in '93, and we did a bronchoscopy on him in '94. He developed severe sputum atypia 14 years later and had a repeat bronchoscopy. This is what his CT looked like at the time of that repeat bronchoscopy. There was an obvious mass lesion here at the superior segment or RB6 by our nomenclature. And then we go back to his previous biopsy, which you see here on your left, showing a moderate dysplasia at this area and the more recent one showing an invasive squamous cell carcinoma.

So if you look at the bronchoscopy results, we usually biopsy at least six sites, and they're standard. You can see that back in '94 he had moderate atypia at this area, and that's where he developed his cancer 14 years later.

So this is, you know, very convincing to some people, but I'd like to also point out that it's very unusual that we ever get a biopsy from a site that later pops up with a cancer. And, in fact, in our bronchoscopy cohort, I believe there are only two patients like this. So this may shape your thinking, but it's very hard to study.

So bronchial histology can be reproducibly rated by pathologists if they are interested in doing so. Many pathologists aren't very interested in this because there's really not a whole lot of clinical impact of this. But it can be reproducibly rated by pathologists, and it goes basically from normal up to invasive cancer, which I'm not showing here, and the grades go from 1 to 8. And so for some of our studies, we just use these numerical grades to, you know, develop averages or whatever for the statistical analysis.

Now, I'm going to spend a little time about the -- what I call the field effect. I think it was originally called field cancerization in a paper by a pathologist named Slaughter, published in 1953, and what Dr. Slaughter noticed was that in pathologic specimens from patients with oral cancer, he could often find multiple preneoplastic lesions and sometimes even other primary invasive cancers, and so he felt that the whole

mucosa had been damaged. And so that has shaped thinking in this field.

So here I have -- on the right-hand side you see that little red starburst, which is a cancer, and the idea is that if this field effect were true, that we would expect biopsies taken at these various areas, denoted by the arrows, to also be more advanced in people with a prevalent cancer, and that perhaps that would help us define dysplasia as a risk biomarker.

Now, the mechanisms for the field effect, I think some of our speakers have already sort of touched on various potential mechanisms. One might just be the intensity of the exposure to a carcinogen. Another might be your genetic susceptibility. We have found that in some cases there will be a progenitor epithelial cell that gets mutated and actually can be found throughout the lung in a minority of the population. And then, of course, epigenetic alterations, as our previous speaker mentioned. So all of these things probably factor into this.

So we've done most of our work with a bronchoscopy cohort that we've been collecting since the early 1990s with our SPORE grant, and about a year ago when we looked at what happened to these people -- we look at it every year, of course -- we had

about 938 subjects. About half of them are dead now. We have lung cancer having either been prevalent or incident in about a third of the patients in this cohort.

And the incident cases surprisingly are not that great. We only have 33. We followed these people through various passive and active methods. So we think we're probably not missing a whole lot of incident cases. And so we can go in and study various -- test various hypotheses using tissues from this.

And this is a case-control study we carried out a number of years ago, where we looked at four groups, and they were healthy never smokers. We had current or ex-smokers with normal lung function and no lung cancer, and current or ex-smokers with airflow obstruction, and current, ex-, or never smokers with lung cancer. And we looked at the histology and the proliferation by Ki-67 index in all the nonmalignant biopsy sites.

And I know this is too busy to read, but I'm going to highlight a couple areas to focus on. These are our results basically. We have in the upper part histology and in the lower, Ki-67.

So gender or sex turned out to be an important determinant

of both histology and Ki-67, even after adjusting for smoking intensity and age. Age was also -- well, let's see, age wasn't that important. Smoking status also obviously was very important, both in terms of unadjusted and after adjusting for other covariates.

Now let's look at the more interesting question of whether this made any difference for disease. And I'll bring that up to highlight those areas.

You can see, initially for lung cancer, if you don't adjust, it looks like histology may be a biomarker. But then when you adjust for smoking and sex, this goes away entirely, which is -- you know, was somewhat of a surprising finding and disappointing, particularly for our pathologists who feel dysplasia is very important.

More recently, though, we're trying to focus in on what -- you know, is there really a signal in this dysplasia? And Dan Merrick, one of our pathologists, has analyzed a subgroup of our bronchoscopy cohort who've had multiple bronchoscopies, and he divided the dysplasias into those that were persistent -- and persistent dysplasia was one that stayed the same across multiple bronchoscopies or got worse -- versus a regressive histology, regressive dysplasia. And a fair number of these

dysplasias will just go away on the next bronchoscopy. Perhaps they're not as, you know, genetically as deeply embedded.

And he found that looking at the incident -- let's see, can I highlight this? Looking just at incident squamous cell cancer, he did come up with a statistically significant signal here with, you know, a fairly wide confidence interval. So it may be that these persistent dysplasias are the ones we should be focusing on as a risk biomarker.

Here in the graph you can look at also more of a prevalence analysis where he found that in cases with invasive squamous cell, there was, you know, a higher frequency of multiple persistent as opposed to regressive bronchial dysplasias.

Now, also you note that I'm just pointing out squamous cell carcinoma. There was no relationship between either of these dysplasias and adenocarcinoma, which I think is also somewhat of a surprising finding, although of course dysplasia is, you know, in the squamous pathway. But you would think that the collateral damage sort of would also give you a relationship with adenocarcinoma, but we didn't find that.

So another way of looking at trying to parse out the dysplasias that are more ominous would be to look at mutations

within those dysplasias. And our initial attempt on this used spectral karyotyping with Marileila Varella-Garcia's expertise. Now spectral karyotyping, or SKYFISH, requires metaphase spreads. So we have bronchial epithelial cell cultures derived from some of these dysplastic areas, and this just shows an example of this. And she divided the abnormalities or the findings up into three. Basically, in the blue, you see normal where all the karyotypes were normal.

Abnormal but non-clonal means there was a single cell with a translocation or a deletion or something like that. Those perhaps could arise from just an in vitro artifact.

And then there was abnormal clonal where there would be multiple colonies showing the same abnormality. These probably are more likely to represent something that was going on in vivo, although they could still arise probably from in vitro artifact as well.

And as you can see, in the never smokers, the majority of the cells were normal, whereas in the patients who had carcinoma -- these are biopsies, of course, taken not from the cancer but from a dysplastic site somewhere else. In the patients with carcinoma, the normals are decreased, and of course the abnormal clonals appear to be risk associated.

I must tell you this is not statistically significant. It's really just exploratory. The numbers were quite small.

So to take this further we looked at FISH, which allows you to actually look at interphase spreads, but you have to focus down on several loci. You can't just do sort of a wide analysis. So FISH in bronchial biopsies we applied to a good sample of our advanced dysplasias. And here's what we found, that here there did appear to be a statistically significant association between chromosomal aneusomy and the presence of cancer at another place in this patient's airways.

And we're not the only people who've worked on this area. A couple of the papers that I think are particularly strong I'm listing here, particularly this one by van Boerdonk. I'd like to just summarize what they did, to let you know what a hard task it is to actually work with biopsies at an area that turns into cancer.

They've done a lot of bronchoscopies on very high-risk people, and in 1,600 subjects, they were able to find six biopsies that they've taken at areas that later became cancer. And they had very painstakingly taken a biopsy for peripherin-embedded formal and fixed analysis and another one that they froze and just kept in a freezer. And so those six biopsies

where they had, by blind luck, found a place that turned into cancer, they looked for abnormalities. They found a chromosome 3q amplification in all six, and of course, you know, in a number of controls didn't find that. So they felt this was indicative that genetic alterations are really a risk biomarker.

Now, we've taken another approach. One of our guest scientists -- Ichiro Nakachi and Chris Coldren have taken an approach where it's unbiased basically. They've utilized high-density SNP arrays to interrogate the DNA of dysplastic sites, and I'll show you basically what they did.

They've taken our biopsies, and we find that if we isolate DNA from the biopsies, you can find somatic chromosomal alterations. But the data is very noisy, of course, because these biopsies contain epithelial cells and stromal cells. So you've got really -- if you're interested in the epithelium, it's really in a minority.

So we've gone more to brushings, and this just shows you that basically you can brush and you get about 80% epithelial cells with a brush. You can vortex and you get a bunch of them off, but a lot of them still stick to the brush, so we've gone more to just working with the whole brush.

Now, if you apply this to clinical samples, it's quite noisy, and people can interpret this, but it's difficult. The upper bar here is the log of the R ratio, which basically measures intensity of the signal at each of the two and a half million loci, and the bottom is the B allele frequency where it's sort of the classic AA, AB, BB, that you find. And I'm not going to go through really how to interpret, but this is pretty daunting to try to pick out the abnormalities, although it can be done.

What they did was they developed -- they also looked at peripheral blood genomic DNA from these individuals and then subtracted that out, which gives you a much nicer picture where they call this the delta-theta, which is really the differential. And here you can really focus on areas that are abnormal.

At the very bottom I'll note that we have the -- here you have the different chromosomes running across. So here you can pick out that there are a lot of abnormalities in this adenocarcinoma.

Now, if you take it to dysplastic biopsies, here's a carcinoma in situ, and you can see, well, there's a 3q abnormality and 3p abnormality and some others. This can allow

you to detect mutations, I think, in a very sensitive way.

Here's a moderate dysplasia brushing, which had a lot of changes. And I'll just basically -- and I should mention that we feel that the sensitivity of this is we can pick up perhaps a population of mutant cells that are 10% of the total epithelial cell population. So it's very dependent on how large the dysplastic area is that you've brushed.

So overall, our results, which have been published, you would -- you know, the numbers here are quite small, but in general, I think you get the idea that for normal or just hyperplasias, it's very unlikely to find somatic chromosomal alterations, where in the dysplastic, sort of the moderate dysplasia or worse, you do occasionally find mutant cells. And in carcinoma in situ, it's quite common. In invasive cancer also it's quite common.

So at this point I'm going to shift gear to another way of looking at the epithelium. We've recently begun to collaborate with Moumita Ghosh, a developmental biologist, and she has really enriched our ability to study the dysplasias, and she isolates airway stem or progenitor cells. People sort of get upset if you call something a stem cell, so we've backed off and called it progenitor instead.

A progenitor cell has two characteristics. It can self-renew, and it can differentiate into multiple cell types of its home tissue. Markers for this are keratin 5 and p63 in the airway epithelium, and the original hypothesis was that the malfunction of these stem cells leads to dysplasia.

Now, this is basically our approach. We do endobronchial biopsies at areas on autofluorescence that look abnormal. We have a bronchial map which is basically where we store our data. And the study design is we have two biopsies from each site. One goes to the pathologist to tell us what it was, and then the other Dr. Ghosh cultures. They come from a couple of our SPORE trials. She's looked at now over 200 of these biopsies.

So the progenitor cells have a characteristic morphology of the colonies. They start from a single cell. She's done limiting dilution to demonstrate this. They display the appropriate markers, and when she looked at the progenitor cell self-renewal, in other words, how many of these colonies of progenitor cells do you get from a thousand basal cells that you put in to culture, she finds this relationship between the grades of dysplasia, basically an inverse relationship. So there are fewer progenitor cells that you can isolate out of

these dysplastic lesions. Interestingly, when she looked at lung function and progenitor cell count, she also found an inverse correlation.

And I'd like to point out that these are for the current smokers. If we add in the former smokers, it's noisier, and I think there's some biological message here, but we haven't figured it out yet.

Now, as far as multipotentiality, this shows you cultures of the progenitor cells stained for intercellular junctions with the zonula occludens and for cilia with the ACT. And I'm getting close to the end of my talk, so I'll speed up here a little bit.

But you can see that, in the dysplastic, the B comes from a dysplastic area, and they don't have the multipotentiality that they do from a normal site.

And this insert here just shows that there are also mucin-secreting cells that come out of the progenitor cell cultures, and they're also reduced out of dysplasia.

Now, we've looked at mutation within these cultures and tried to correlate it with the airway brushings, and basically we've found, using RNA sequencing, you know, common variants that we think are probably mutation, and we've actually, in one

case now, verified these mutations by whole exome sequencing. So we've found this on the RNA level and verified it on the DNA level, comparing to peripheral blood, showing that it's a somatic cell mutation.

So finally, I'm going to finish up with a quick run through chemoprevention, and this is a summary of Phase III chemoprevention trials. You'll notice they've all been negative or have generated harm. I think everyone in the audience probably knows this.

Intermediate endpoint trials have been pretty disappointing as well, including a number of trials with dysplasia as the outcome. There have been some trials with celecoxib that have been positive for decreasing Ki-67, so I don't list those.

Now, our group has been interested in the prostaglandin pathway that Dr. Hecht mentioned, and as you know, this can be suppressed upstream by various compounds. Downstream, however, you can see there are a variety of compounds, some of which potentially could be beneficial, others of which, like PGE_2, are probably not. Iloprost is a PGI_2 analogue that is somewhat more stable than PGI_2 itself.

We've made prostacyclin synthase over expresser mice.

They're very protected against cancer, against lung cancer by a variety of chemical carcinogenesis approaches.

And we've done an iloprost chemoprevention trial, basically with two bronchoscopies and placebo or iloprost in between. And this is the result, which I'll run through quickly.

We used average histology score as our primary endpoint. And as you can see, if you looked at all completers, this did not achieve statistical significance, although it might have if we've used worst. But then -- you cut me off. No, I cut myself off. However, if you look at former smokers, everything got better. And if you look at current smokers, nothing got better.

So if you look at the response proportion, about half of the former smokers improved with iloprost. And when we go back to the progenitor cell cultures, indeed very nicely, they also respond in some cases to iloprost, as shown here, but not in all. So it's about half a response there, but we haven't yet matched that up with response in the trial.

Thank you for your attention. And this lists the contributors and our funding.

(Applause.)

DR. DRESLER: Thank you, Dr. Miller.

Let's go ahead and take the break now. And let's take a 10-minute break, and when we come back, then we will hear from Dr. Spira, and then we'll do the panel and then break for lunch. We have this time because we didn't have that full hour for the public speakers. So it is now just before. Let's come back at 20 of, so 10:40, all right, and we'll start again.

(Off the record at 10:28 a.m.)

(On the record at 10:42 a.m.)

DR. DRESLER: So the plan is we'll have one presentation, then we'll do the panel, and then we'll break for lunch, okay? So lunch will be a tad early.

So our next speaker is Dr. Avrum Spira from Boston University School of Medicine. He will be speaking on The Airway Transcriptome as a Biomarker of Tobacco Smoke Exposure and Lung Cancer Detection.

DR. SPIRA: Well, thank you very much for including me in this very unique workshop. I apologize, I wasn't able to be here yesterday, but I'm glad I'm able to be here today. And actually, I think, very appropriately my talk follows York's in that a lot of the concepts embedded in his talk will be expanded upon in mine. So I'm going to focus, though, on the

airway transcriptome as a molecular biomarker of tobacco smoke exposure and lung cancer risk.

Before I do that, though, let me get my disclosures out of the way. I am a founder and consultant to a molecular diagnostic company called Allegro Diagnostics. That was acquired by a publicly traded company about a year and a half ago. Actually, I know it was September 4th, 2014. My wife likes to call it my second birthday. So that is a company that has commercialized one of the biomarkers I'll be showing you today.

So there are three types of biomarkers that I want to walk through with you today over the next 25 minutes. The first at the top is a mature bronchial airway gene expression biomarker for detecting lung cancer early. This test has moved from bench to bedside, and it was launched commercially about 9 months ago, and I'm going to spend a good chunk of my talk on that biomarker. I'm then going to move on to our efforts to go from the bronchial airway to the nasal epithelium as a less invasive biomarker for detecting lung cancer. This is earlier work, not quite ready for prime time, but I want to present to you the potential of that approach.

I'm then going to end by talking about moving beyond

tobacco smoke and using airway gene expression as a biomarker of other inhaled toxic exposures, including electronic cigarettes as well as biomass combustion, which of course is a growing problem in the developing world.

Everything I'm going to tell you about today is predicated on a relatively straightforward concept called the field of injury. And the idea here is if you smoke or you're exposed to other inhaled toxins in the environment, all of the epithelial cells that line your respiratory tract from nose and mouth through the bronchial airway over here have alterations in the genome of the epithelial cells due to the exposure into the toxins in cigarette smoke. Now, we believe that there's variability between smokers for how your epithelial cell system responds to and is damaged at a genomic level from that exposure, and that variability in the gene expression response in these cells associates with your risk for developing lung cancer.

I spent part of my day yesterday at the NHLBI showing that this same concept is true for COPD, where we've developed molecular biomarkers in the airway for that disease. And, in fact, one of the NHLBI program officers was excited about exploring the idea of the field of injury being a biomarker for

cardiovascular disease from smoking. So again, I think this concept goes potentially well beyond lung disease.

Now, based on that idea, our lab has spent much of the last decade developing gene expression-based biomarkers in the bronchial airway epithelium for measuring the physiological response to smoking and a risk for lung cancer. All of the studies on this slide were done in cells that we obtained at the time of bronchoscopy. And you heard a lot about that from York.

But this is the scope being introduced via the trachea, over here, just after it bifurcates into a right and left main stem bronchus, so we're not going very far in. And we use an endoscopic cytobrush through the working channel of the scope, shown over here, to rub up against the inside of the airway wall at that site. So we're not biopsying the airway. Similar to what you saw in York's presentation, we're brushing the airway, giving us a relatively pure population of epithelial cells.

At the very end of the talk, if I have time, I'll show you that this population, in fact, is more heterogeneous than we think. And when we do single-cell RNA-seq -- and I'll show you a little bit of that data at the end -- we do see quite a bit

of heterogeneity between epithelial cells in the airway for the response to smoking.

Nonetheless, we used this approach initially back in 2004 and 2005 to look at healthy smokers and nonsmokers, and we found that active exposure to tobacco smoke alters the activity of hundreds and hundreds of genes in your bronchial airway epithelium. No big surprises there.

What was a bit surprising was in 2007, when we went to former smokers, we found that most of those changes revert back to baseline within a few months of quitting, but about 20% of the alterations persist decades after you quit smoking. So there's an irreversible persistent signature of tobacco smoke damage in these cells.

And that led to the third and clinically relevant marker on this slide, which is a gene expression signature in these cells that could distinguish both current and former smokers with lung cancer from those with benign disease of the chest. And this is a biomarker that's now moved into the clinic, and I'll show you more about that in a moment.

I just want to mention, at the bottom of the slide, that we believe the same paradigm can be applied to COPD. And in the past 3 years, we've developed a series of gene expression

biomarkers in these same cells that can be used as a companion diagnostic for that disease, specifically who responds to inhaled corticosteroids.

Now, before I show you the cancer marker, I wanted to share with you some unpublished data that we just generated with a group in Europe that I am very excited about. This is Dirkje Postma and Maarten van den Berge in the Netherlands.

This is a question I often get at all of my talks: How little do you have to smoke to actually see changes in airway gene expression? It's a hard study to do because you got to take a nonsmoker, get them to start smoking, and bronchoscopy them over time. Tough to do here. Not so hard to do in the Netherlands apparently. So what they did there, they took party smokers, young folks that smoke a couple of cigarettes on the weekend, essentially nonsmokers, had them abstain from smoking for at least a month, did a bronchoscopy, gave them three cigarettes a couple of hours later to smoke, and then re-bronchoscopied them 24 hours after that. It's an incredible study if you think about that design. And York appreciates it.

So we then did gene expression pre and post. A really cool study, right? And what we found is as little as three cigarettes alters airway gene expression. This is a set of

about a hundred genes that change in expression within an individual who's exposed to those three cigarettes.

And I think what's particularly intriguing is when we look at this list and compare it to our prior publications, cross-sectional studies of older chronic smokers, we see a lot of overlap between the gene lists. That is to say, a lot of the genes that we saw in the older chronic smokers appear to change with just a couple of cigarettes or three cigarettes after 24 hours.

But I want to spend most of the time today talking about lung cancer and our biomarker, and let me begin. Hopefully, this will -- there we go. I'm going to show you a small video. One of the advantages of starting a company and raising venture capital financing is you learn attention spans are very short. So I'm going to try to explain our biomarker in 45 seconds or less, and then you can kind of shut your eyes and doze off for the rest of the presentation. So sort of like the venture capital community.

So here we go. So here's how the test works. So when you smoke, every time you take a breath in, you're inhaling all of those toxins and affecting all of the epithelial cells that line your respiratory tract, and over many years of that

exposure, a tumor can begin to grow deep in your lung over here. And if you're fortunate enough to have a CAT scan of your chest done at this point in time, either incidentally or for screening purposes, you might catch the tumor when it's, let's say, over here, 1 cm in size. The problem is, on the CAT scan, we can't distinguish a benign from malignant nodule.

So often we'll do a procedure like bronchoscopy, shown here, to try and make a diagnosis. The problem with bronchoscopy in the end of the day is the scope cannot get all the way out to where this nodule is. The diameter of the scope precludes easy access to this compartment. So often the bronchoscopy will be non-diagnostic, and you have a clinical dilemma what to do next. Which of these patients do you biopsy the lesion on, either a needle through the chest wall or a surgical lung biopsy, versus who can you afford to watch the lesion with repeat imaging of the chest and only biopsy if the lesion grows? And so that's essentially the unmet need that we were trying to address.

So what we told -- sorry, that should not have gone back. Hold on. You're going to see the video again. Let's see, if I go to the next slide, hopefully -- here we go.

So what we asked the pulmonologist to do is do all the

routine studies at the time of bronchoscopy, but before they pull the scope out of the airway, spend an extra 30 seconds and brush the normal-appearing mucosa in the right main stem bronchus. We take those cells, run them on a microarray, and the pattern of gene expression tells us whether the nodule deep in the lung is benign or malignant. So in the event that the routine cytology is non-diagnostic, you can use the gene expression information to decide biopsy that lesion right away or I can afford to watch it with repeat imaging. And that's essentially our test.

Now, as I learned, the *New England Journal of Medicine* doesn't like videos. They actually like data, so let me try to show you a couple of data points to argue that this test really works.

We did two clinical trials to develop and validate the marker. They're called AEGIS I and II. They were funded by Allegro Diagnostics. It involved 28 medical centers shown on this map, mostly here in the U.S., although we had one medical center in Canada and one in Europe.

Here's a summary of the study, and I'll just show you a couple of points from the *New England Journal* paper. It involved a thousand or so current and former smokers undergoing

bronchoscopy for clinical suspicion of lung cancer, and they were enrolled into either AEGIS I or AEGIS II depending on the time point of enrollment.

Now, we took half of the 600 patients in AEGIS I and used them as a training set to refine and develop a 23-gene biomarker. That biomarker was then locked down, published, and validated subsequently on the remainder of the samples in AEGIS I and all of the AEGIS II test set samples.

So importantly, we're measuring these 23 genes in the bronchial -- the normal-appearing bronchial epithelium of the main stem bronchus. So we're asking the pulmonologist to do all the routine biopsies and brushings, but before they leave that airway, just take an extra brush from the main stem bronchus. That brush is sent to a lab in North Carolina where the 23-gene biomarker is measured.

We then need to follow these patients after their bronchoscopy until a final diagnosis is made. It's a prospective study. So we follow them for up to 1 year so we know whether it's cancer or not, and then we can look to see whether the biomarker was accurately predicting patient classification.

And here's the bottom line, and I'll show you the data in

a moment. The overall sensitivity of our biomarker was 90%, which is good news. The bad news is the specificity is modest. It was 50% in the study.

But the negative predictive value, as I'll show you, was above 90%, and the NPV is what will drive clinical utility in this setting because physicians want to essentially rule out lung cancer. They want to know who can we avoid sending for an unnecessary invasive procedure.

And so based off of that data, Veracyte launched a CLIA version of this test, called Percepta, almost a year ago now, actually. It's been launched as a registry study in 25 medical centers here in the U.S. The test is being offered for free at this point, and they're really trying to gain data on real-world use of the test, which I think is appropriate before wide-scale pushing the test out into the clinic.

I'm just going to show you a couple of quick data points from that *New England Journal* paper.

The first is shown over here with Table 2, where we stratified patients based on size and location of the lesion on CT. So our study are all comers to bronchoscopy for clinical suspicion of lung cancer. About 210 of those patients had a nodule that was under 3 cm in size, as shown here.

I'm showing you in this column the sensitivity of bronchoscopy alone for detecting lung cancer in that size range. It only picks up about half of the lung cancers, whereas the classifier, the gene expression test, picks up more than 90% of the lung cancers that are under 3 cm in size. And most importantly, if you combine our classifier with the results of bronchoscopy, which is how we see the test working in the real world, you've got more than 95% sensitivity for detecting lung cancer.

Sorry, I'm going back here. I just want to mention, at the bottom of slide, we're stratifying the lesion by location on the CT. About 194 patients had a peripheral lesion on their CAT scan. Again, bronchoscopy sensitivity is modest, only about 55% for peripheral lung cancers, whereas the biomarker has 90% sensitivity and combines 95% sensitivity for picking up peripheral lung cancer.

So this is really a sweet spot for the test, small peripheral lesions where the bronchoscopist has the hardest time accessing the lesion.

The other data point I want to show you from the trial is Table 3 of the paper, where we stratified patients by pretest risk of disease. This is very important because you can

imagine if you have very high pretest risk of disease, there's no biomarker in the world that's going to change what a pulmonologist does next after a non-diagnostic bronchoscopy. They're going to stick a needle into the thing. So what we asked pulmonologists to do here in the study is define the pretest risk of that patient prior to their bronchoscopy and stratify it into low, which we defined as under 10%, intermediate, which was 10% to 60%, and high or greater than 60%. You can see we had 101 patients that were in that intermediate pretest risk group. Bronchoscopy did not do particularly well in that subgroup. Its sensitivity is only 41% for diagnosing lung cancer, whereas the classifier of the gene expression sensitivity is high and specificity is just around 50%.

And the most important number is over here. The negative predictive value of our classifier is 91%, meaning if you have an intermediate pretest-risk patient, a patient where you're on the fence, is this cancer or not, and you do the bronchoscopy and it's non-diagnostic, if the gene expression test is negative, there's a less than 10% chance that patient has lung cancer. And that's someone you might follow then in the clinic with repeat imaging as opposed to biopsying.

And so this was -- that's really the clinical utility of our test. Demonstrating clinical utility, though, is a whole other thing. We learned it's a huge challenge to overcome the classic valley of death. Clinical validation is great, but showing you're going to impact patient care is very, very challenging. In particular, in an observational trial where we didn't give the results back to the physician -- so physicians, they couldn't assess how this would change their management.

But one of the things that we did do is calculate the negative likelihood ratio for our test, which is another metric for converting a pretest risk of disease to a post-test risk. And it looks like a hockey stick over here. I grew up in Canada, so everything looks like a hockey stick.

But this is basically how it works. In our test, if you have a pretest risk of up to 60%, that would convert your -- a negative classifier score would convert your post-test risk to 10% or lower. So that's where we think this test can have clinical impact.

And one of the other things we've now done, and this is work that's about to be -- actually, it's been accepted for publication and will be coming out shortly -- is we modeled clinical utility of our test in the AEGIS trial. We looked to

see what happened to patients post a non-diagnostic bronchoscopy, and it turns out almost half of the people whose bronchoscopy was non-diagnostic and who turned out not to have lung cancer had an invasive procedure done unnecessarily after the bronchoscopy, either a needle through their chest wall or a surgical lung biopsy.

And this is the number that's truly scary. One in three patients -- sorry. One in three patients who had a surgical lung biopsy in our trial had it for benign lung disease. And that's consistent with what we see across the U.S. for benign thoracotomy rates. And the reason for this is we never want to miss a lung cancer, but think of all the morbidity and actually mortality associated with a surgical lung biopsy.

So the idea here is if we have 50% specificity, which is what our test has, we could avoid about half of those invasive procedures that are unnecessary, assuming physicians acted on the results of our test.

Now, I'm going to skip over this. Let me move into sort of moving into the nasal epithelium. So this test is great if you're undergoing bronchoscopy. And the slide I skipped over showed you there about 100,000 people a year in this country who have a non-diagnostic bronchoscopy for suspicion of cancer. So there is a clinical unmet need, but we recognize if we could

move out of the bronchoscopy setting to a less invasive setting, this test could be applied more widely.

So we were very excited a couple of years ago to begin to move into the nasal epithelium as a less invasive site. Let me try to convince you it's less invasive. Folks often recoil at the idea of sampling the nose. What we're doing here is brushing the inferior turbinate.

So if you press on your nostrils and look up into the mirror, you'll see a triangular protuberance on the lateral wall of your nose, over here. That's the base of the inferior turbinate. It's about a half a centimeter in. We just take a little endoscopic cytobrush over here, brush up against the base of that turbinate for 2 or 3 seconds, and collect a huge amount of high-quality RNA. It's actually remarkable. No local anesthesia, extremely well tolerated. We've done this now on hundreds and hundreds of adults and hundreds of children, as I'm about to show you. Again, it is noninvasive.

So initially what we did with this approach is demonstrated that if you take someone who's actively smoking and a group of people who aren't and do a bronchoscopy and sample their nose and their bronchus on the same visit and parallel run them on a microarray, we find that virtually all

the changes with smoking in the bronchial airway are mirrored in the nasal epithelium within an individual. So there's a high level of surrogacy of the nasal gene expression response to smoking to that seen in the bronchial airway.

And so that allowed us to do a couple of really cool things. You can imagine, now that we have the nose working for smoking, we could team up with Dorothy and Steve, for example, and begin to look at longitudinal sampling of the nose, something we could never do, except maybe in the Netherlands, in the bronchial airway, but do this over time.

And so one of the studies that we were fortunate enough to collaborate on with Steve and Dorothy was a group of patients in their clinic who were undergoing smoking cessation, where they were able to sample the nose pre-quitting and at 4, 8, 16, and 24 weeks after quitting. And then we can define the gene expression kinetics that occur within 6 months of quitting.

And this is just a heat map showing you some of the results that we're hopefully going to submit for publication shortly, a set of a hundred or so genes that rapidly change within 8 weeks of quitting. You can see a group of genes here that are low at the beginning and get turned on within 8 weeks, and a larger group of genes that are turned off within 8 weeks

of quitting. And notably, there's a small group of genes over here that actually change within 4 weeks, a much more rapid response to smoking cessation. This is an example of the kind of information a nasal biomarker approach can yield.

A second idea that we were very excited to pursue with the nose is to go into children, where we can't easily do bronchoscopy, and look and see whether the nasal epithelium could be a biomarker of secondhand smoke exposure, which of course has a lot of effects in children, health effects.

And this is a collaboration with a group at Columbia and another group at Johns Hopkins. They have cohorts of inner-city asthma children who are exposed to secondhand smoke in the home, where they have objective measures in the home of how much secondhand smoke the kids are being exposed to. And we were fortunate to collaborate with them where they could collect nasal brushings from their kids in their home, both a group of kids that had been exposed to secondhand smoke and a group that had not, as well as a group of adults. What I'm showing you here is a gene expression signature that we developed that can distinguish secondhand smoking adults and children from those that are not exposed to secondhand smoke. So it's a very sensitive biomarker of low levels of tobacco

smoke exposure.

We've now created a biomarker from this gene signature and are beginning to test it on independent cohorts to see how well it can predict secondhand exposure.

But what I want to end on is we want to now move the nasal test into the lung cancer diagnostic setting. So to do that, we actually leveraged the AEGIS trial, that trial where we validated our bronchial biomarker. Every patient in that study had a nasal swab done that was run on a microarray in parallel. And what we were able to show, and I'm not showing you it here on this slide, is that there's surrogacy of the bronchial gene expression changes for cancer in the nasal epithelium. We see a lot of overlap within an individual.

And we went ahead and developed about a 200-gene signature, shown here, that distinguishes the nasal epithelium of patients with lung cancer, in red, from those without lung cancer, in black. It's not perfect, but it's a pretty good separation. We then developed a 30-gene biomarker from this 200-gene signature and tested it on an independent set, and it should read over here, of 130 patients in the AEGIS II trial.

And I'm showing you here the performance of our genomic biomarker compared to clinical risk models that predict lung

cancer in this setting. In particular here, the Gould model, which looks at size of the lesion, smoking, and age, it actually does pretty well, as you can see, at predicting lung cancer in the AEGIS trial. But if you add the 30 genes in the nose to the clinical model, you improve the AUC. It's a modest improvement, but it's statistically significant. And more importantly, you get about a 10-point increase in sensitivity and a 15-point increase in NPV. And as I told you, those are the key clinical parameters for impacting care in this setting.

Now, that's, of course, the setting of patients undergoing bronchoscopy. That's not where we see the nasal test having its best clinical utility. There are two clinical settings that we currently are working on. One is post a chest CT where you have a small nodule for which the patient's not going to undergo bronchoscopy, where we think a nasal diagnostic test would have utility. But I would argue, the more exciting is upstream of chest CT. So imagine smokers, at their annual physician's visit, have a nasal swab done, and based on their gene expression profile, you stratify those patients into undergoing chest CT screening, which of course we don't know who to pick for right now in terms of best candidates. So that's what we'd like to do.

One of the challenges, though, with this biomarker or developing that marker is to have data to support that nasal gene expression changes years before you get lung cancer. It's a very hard study to do.

One of the ways we're doing that now is through a consortium called DECAMP, that some of you, including York, is involved with in the room. It's a $20 million grant we got from the Department of Defense 4 years ago, allowing us to bring together leaders in the molecular biology world and lung cancer, as well as a number of cohorts from military and VA hospitals across the country.

And the bottom line. There are two clinical trials that we're actively involved with right now. I want to focus on the one, the second trial on the bottom, for this talk. This is a screening study of 800 high-risk smokers who either have COPD or a first-degree relative with lung cancer being followed for 5 years, with annual CTs of the chest as a screening tool and every biospecimen imaginable being collected annually, including bronchoscopies being collected at two different time points. So a very aggressive biospecimen collection protocol. And the idea here is can we develop molecular markers, including in the nose and airway, that predict future risk of

cancer?

Let me end by talking in the last 2 minutes here about moving beyond tobacco smoke. So we were fortunate a couple of years ago now to be funded by the FDA and the NCI through one of our EDRN, or Early Detection Research Network grants at the NCI, to basically take this gene expression approach in the airway for smoking and bring it to electronic cigarette smoke.

And we call this the ECAPS study. It's still under way, but we basically have an in vivo arm where we're developing bronchial and nasal gene expression biomarkers in e-cig users, as well as an in vitro arm to sort of model the effects in a culture system.

I'm just going to share with you a couple of quick snippets. This is preliminary and still in progress. This is the in vivo data from actual patients. It wasn't as easy as I thought to recruit the cohort, to get e-cig users to undergo bronchoscopy as volunteers.

And we matched. These are former smokers who used e-cigarettes currently, for the last 6 months at a minimum, matched to a group of former smokers who do not use e-cigs. We see a very strong -- this is an *n* basically of 10 in each group, and we see a very strong signal for gene expression in

the bronchial airway of e-cig users.

And what's particularly interesting about it is many of these genes overlap with genes that we see changing in the bronchial airway of regular tobacco smoke users, arguing there are some similarities in the airway gene expression response to e-cigarettes and tobacco smoke, which I think is obviously concerning.

Now, again *n* of 10, early days. We're just recruiting additional patients into this trial now, but we've actually done the in vitro experiments in parallel that support these in vivo findings. So this is an air-liquid interface system where we take bronchial brushings from a nonsmoker, grow the cells on an air-liquid interface for 21 days so they differentiate into all the cell types we see in vivo, and then use whole vapor e-cig or traditional cigarette smoke and expose the cells and then do gene expression profiling.

And what we find -- and I won't have time to go through the details here, but at the very, very top, when you do a principal component analysis of all of the genes that change, you find that e-cigs -- the e-cig exposure lies somewhere between t-cigs and air controls. So there are some genes that mimic what we see with traditional tobacco smoke, but some

differences. So it's somewhere in between. Perhaps not a big surprise.

What was interesting is -- and this is something we can't do in vivo, but only in vitro, is modulate what's in the e-cigarette and see what is inducing the gene expression changes. And just to show you a snippet of that, it turns out tobacco flavoring and the presence of nicotine in an e-blu cigarette are what induce most of the gene expression changes. Those are the components that we think are responsible. Again, early days, but that's what it looks like.

The last data slide I'm going to show you is a second example of moving beyond tobacco smoke, a study that we just published in *Carcinogenesis* with Nat Rothman and Qing Lan at the NCI. They've been following a really unique cohort of women who are nonsmokers in China but have very high rates of lung cancer by virtue of being exposed to indoor air pollution. They heat and cook in poorly ventilated homes and are exposed to smoky coal fumes 24/7.

And what Nat and Qing were able to do is -- we couldn't even get nose, we couldn't convince Nat and Qing to do the nose here. We certainly couldn't get a bronchoscopy. So we had buccal. These are swabs of the cheek, which the RNA is

partially degraded, but nonetheless were able to be put on microarrays and develop a gene expression signature that correlated with the degree of exposure in the homes that Nat and Qing are able to measure in terms of benzopyrene exposure.

And what's interesting again is these genes overlap with what we see in cohorts in Boston in the buccal epithelium with regular tobacco smoke exposure, so again arguing that we may be seeing something nonspecific in terms of an airway gene expression response to inhale toxins. And again, this is all in healthy individuals. We're now beginning to work with Nat and Qing in women who do or do not develop lung cancer with exposure.

So let me end there and just simply summarize what I've shown you, which is on this slide. Bronchial airway gene expression, we believe, is a good measure of the physiological response to smoking and potentially can serve as a biomarker for identifying smokers with lung cancer. And, in fact, the first test, commercial version of the test has been made available as of 9 months ago.

We think the same paradigm extends to the nasal epithelium, which would provide a less invasive site to perhaps develop biomarkers not only of tobacco smoke exposure but of

lung cancer risk.

And finally, we think this approach potentially could serve as a new paradigm for rapidly evaluating additional products like electronic cigarettes, in terms of their effects on airway gene expression.

And oh, I forgot I had this one. This slide, this is the last slide here, just to make the point that everything I've shown you until now are brushings of the airway, about 100,000 cells profiled together.

Genomic technology has evolved at an unbelievable pace in the last 10 years. We now have the capacity to do RNA expression on a single epithelial cell from the bronchial airway.

So this is work we'll be presenting at the ATS meeting next month, from Grant Duclos, Jennifer Beane, and Josh Campbell in our group, where we took those same brushings and sorted them down to 1,000 single epithelial cells from each patient and did RNA sequencing of each of those individual 1,000 epithelial cells and did this across a group of smokers -- it's crazy stuff, but true -- across a group of smokers and nonsmokers.

And what's really cool about it, and I don't have time to

go through the data now, is you can see heterogeneity in that gene expression response to smoking. There are subpopulations of cells in that bronchial epithelium that are producing most of the signal in terms of the response to smoking. And it's really, really cool to begin to dissect what those subpopulations of cells are and what that might mean for lung cancer risk, and that's something we're just beginning to do.

So I'll end there. I just want to acknowledge a lot of collaborators outside of BU, many of them -- some of them in the room today, and of course all the individuals at Boston University involved with this work; and finally our funding, including the FDA, for some of what I've shown you today.

So thank you.

(Applause.)

DR. DRESLER: And the other speakers, if you could come on up for the panel, please.

Okay, so we'll do the same as we did yesterday. If you have any questions, you're welcome to come to the microphone. If you're online, please send in your questions, and those will make its way up. And if you don't want to go up to the microphone, just raise your hand and you'll get a card, and then we'll get the questions up here. So I'll start out with

some questions for you.

Steve, one of the questions that I had from your presentation was why are women more likely to have 8-epi-PGF? Why did they have higher levels?

DR. HECHT: The answer is I don't really know. But I mean, that's what our data shows us so far, but it's pretty limited. So I think also, in the Total Exposure Study, they saw higher 8-epi in females than males. Is that right?

DR. FROST-PINEDA: I know BMI was the most important factor, and then nicotine equivalents was the second most important factor. So I'm not sure where gender was.

DR. HECHT: Okay. So anyhow, that's what we've seen in our data so far, but it's kind of limited. Only probably a couple hundred subjects. I don't know why.

DR. DRESLER: And then the other question that I had -- and I've always had this clinically, too, in speaking with clinicians. You know, the Japanese smoke different cigarettes than other parts, certainly of Europe or the U.S. So could that be part of the reasons for their differential risk for lung cancer, besides what you were finding?

DR. HECHT: That study, the Japanese Americans are all from Hawaii.

DR. DRESLER: Japanese Americans. I'm sorry. Okay, sorry about that. I missed that.

Okay. Sample collection from airway epithelia histology often requires a biopsy. So we heard yesterday about nasal swabs, and Dr. Tarran -- I think that you're still going to see Dr. Tarran. And I don't know if you had heard that yesterday. You missed that, yes, but he was speaking also about nasal biopsies. So it looks like the field is moving towards less invasive.

DR. SPIRA: Can you guys hear me? Here we go. No, unfortunately I missed the talk. I saw the agenda. That's one of the talks I really wanted to see, but I was down at the NIH yesterday.

Yeah, I think the nasal epithelium -- yeah, we're looking, just to be clear, at brushings. It does appear to be an effective surrogate for the gene expression response in the bronchial airway, and we do see -- as I said, we have unpublished data now on lung cancer, suggesting our signature from the bronchus could move up to the nose.

We haven't looked at biopsies, to be frank. And again, I didn't see the presentation yesterday, and maybe the speaker can comment on whether they're looking at brushings versus

biopsies and whether the cell type differences between a brushing and biopsy may change what you see on gene expression.

DR. TARRAN: I can't comment on whether the gene expression profile is different between brushings and biopsies, but I'd predict they'd be broadly similar. It's just a case of number and you get more cells with the biopsies. But I think both are good. And also, just since you missed the talk, we're looking at the physiology, doing the electrophysiology on the smoke response. It's very similar between the nose and the lower airways, which is nice.

DR. SPIRA: Which is what we see, okay. No, that's nice to see, sort of replication of that finding. We haven't done the biopsies. The concern, especially moving to kids, is how invasive can we get in the nose? What's tolerable without any kind of local anesthesia? And as I said, the brushings we do seem to be -- I don't know what your experience has been with the more aggressive biopsy approach.

DR. TARRAN: Having been biopsied extensively --

DR. SPIRA: Okay.

DR. TARRAN: -- I think definitely for younger people, brushings are probably preferable.

DR. SPIRA: Yeah.

DR. DRESLER: Dorothy, did you want to --

DR. HATSUKAMI: Sure. Dr. Hecht had shown a PowerPoint, our results from a study showing that total NNAL does predict the risk of lung cancer. And I guess, Avi, I was wondering, do you -- have you thought about looking at the correlation between total NNAL levels and gene expression rates?

DR. SPIRA: We would love to. I mean, it's one of the reasons we reached out and started collaborating with you and Steve 5 years ago, was the concept that -- well, I think airway gene expression is a good marker. Obviously, it's got its limitations, and I think it would be very interesting to correlate it with some of the, you know, well-defined markers that your group and Steve have developed over many, many years.

I would anticipate there would be a relationship. I do think how your airway epithelium responds probably relates to how you metabolize the ingredients, the components of tobacco, and probably associates with other markers of risk like NNAL. Again, we haven't done those studies. Perhaps that group or some of the people that we've profiled together with you, if the NNAL has been measured, we can begin to do that. I don't think we've ever done that, though.

DR. HATSUKAMI: Thank you. Can I ask a few more

questions?

DR. DRESLER: Yeah.

DR. HATSUKAMI: All right. Dr. Shiels, you had mentioned that nicotine is immunosuppressive, and I guess I'm wondering, what studies looking -- have been done, looking at nicotine specifically that might be immunosuppressive? And I had that on one of your slides.

DR. SHIELS: So I don't know the studies intimately. I'm happy to provide you with the literature that we cited in the paper. Yeah, I don't know the literature specifically on that in depth.

DR. HATSUKAMI: Sure.

DR. SHIELS: Yeah.

DR. HATSUKAMI: Okay, that's fine. The other question I have for you is that you said that some biomarkers were not responsive to time to quit.

DR. SHIELS: Right.

DR. HATSUKAMI: And so with those biomarkers, if there's some replication of those findings, do you think that if those biomarkers are insensitive, would they be appropriate to use for evaluating tobacco products?

DR. SHIELS: So I would say that the key to what you said

is if they're replicated. So the markers that we examined in the studies that I presented today, a lot of them have not been examined in epidemiologic studies, you know, in humans and circulation with relation to tobacco products. So I wouldn't be comfortable making any recommendations about them being biomarkers of harm without certainly replication. The assays that we used were great in that we were able to measure a lot of markers in a relatively small amount of specimen. But, you know, a lot of these markers have not been really well examined in blood samples.

And so I think the first step is that these results need to be replicated and using a study design that may be a bit more appropriate than the one that we were able to sort of piece together from what we had.

DR. HATSUKAMI: Okay. May I ask one more question?

DR. DRESLER: Sure.

DR. HATSUKAMI: All right. So Steve, I'm going to ask you this question that I had asked before, because you had mentioned that biomarkers -- what we consider biomarkers of exposure are also biomarkers of harm. So in what situations do you think it's appropriate just to use biomarkers of exposure if biomarkers of potential harm are not available?

DR. HECHT: I think total nicotine equivalents emerges as a very useful biomarker, not only of exposure but potentially also as a biomarker of harm, because total nicotine equivalents not only gives you virtually the entire nicotine dose but just about everything that comes along with it. So there's a lot of data that indicates a correlation of various biomarkers with all the mercapturic acids. You look at NNAL. They all correlate with total nicotine equivalents.

So total nicotine equivalents is telling you what the guy is taking in and everything. So I mean, I think that it's readily measured, and it should be taken as a biomarker of harm, investigated as a biomarker of harm.

DR. PRASAD: A truly fascinating session. I truly enjoyed the talks. I have several questions, if I may.

Dr. Hecht, we talked about total -- the Shanghai Cohort Study you presented previously, and then you talked about NNAL biomarker of exposure as it has a potential to become a biomarker of effect. Can you comment on other nitrosamines like NNN? Do you think that will have a potential to be one of that?

DR. HECHT: So NNN we've also looked at. I didn't mention it here, but we looked at NNN with respect to esophageal cancer

in the Shanghai Cohort Study using basically the same approach that I described for the lung cancer, except it was smaller because there were far fewer esophageal cancer cases. But the effect was far stronger than we saw for NNAL, for phenanthrene tetraol, or for total nicotine equivalents. So NNN was related to esophageal cancer, strongly so in that study, but it needs to be confirmed because the numbers were fairly small.

DR. PRASAD: So if I may ask a question to Dr. Shiels. You looked at a panel of cytokines, a large number, and this is pretty much done on Luminex technology. Is there a need to verify these, confirm these cytokine levels by an independent technology? Do you see that need?

DR. SHIELS: I'm sorry, by -- oh, by a different technology?

DR. PRASAD: Yeah.

DR. SHIELS: I mean, I certainly think that they need to be replicated for -- so the results that I presented are from an initial set of three case-control studies that we had carried out, and since then we've sort of, you know, thrown out some of the bad markers and replaced with a better panel, that sort of thing. So the technology is not perfect. For many of these markers, for most of them there isn't a clinically

acceptable, you know, assay like CRP, where you would have a gold standard to measure against. So it would certainly be good to have replication, not just using this technology but perhaps using, you know, standalone assays or other assays with good performance characteristics and compare them with what we found.

DR. PRASAD: Two more questions.

DR. DRESLER: Sure, go for it.

DR. PRASAD: The methylation is truly fascinating, and you talked much about AHRR methylation and correlation. So are there any other markers you can speak, other than the AHRR loci methylation? Any other genes that -- any of the gene loci that come as strongly in smoker versus nonsmoker type of situations?

And also the question for Dr. Spira is you talked pretty much on gene expression. I was wondering if you had a chance to look at methylation to see sort of a predictor. Is there any concordance on that?

DR. BELL: So the first question is there are many CpG locations that are reproducibly altered by the exposure, tobacco smoke exposure. The magnitude of the effect for the AHRR gene, it seems to be the greatest that we can currently measure. But the deficiencies in the sampling of CpGs, we

think, is perhaps one of the reasons why this one is so strong and others are less, because the arrays that have been used only sample about 5% of the genome. And as I showed, we're trying to use additional methods of sequencing to identify if there are other changes that are similar to this gene.

And to speak sort of briefly about the correlation between expression and methylation, this is actually -- Avi and I have been sort of trying to put together a study for several years that just never has quite come together. And, you know, I think we would probably believe that some of the changes, at least some of the changes that are observed would be related to methylation changes that would be somewhat persistent and then could eventually go away. So methylations can certainly be reversible, but as we see, it can be very persistent as well.

DR. PRASAD: I have one other question.

DR. SPIRA: No, it's okay. No, let me add to what Doug was saying. We've thought about those experiments. Our group is obviously a very transcriptome-centered group. We like the transcriptome as a readout because whether it's a microRNA or a methylation of N or a polymorphism, it doesn't affect transcription, and it's probably not important biologically. And so we see this as a common functional readout.

Having said that, mechanistically, it's important to understand why are there 20% of the genes that are irreversibly altered? And I do think methylation is a good process to look at in terms of that question. As Doug said, we've been thinking about studies. I think the methylation platforms, the global, you know, whole genome methylation platforms are not as mature and as robust as gene expression, so that's made it technically a little more challenging.

We have looked, I should mention, with Doug, at polymorphisms in the promoter region of some of the genes in our signature, specifically genes in the antioxidant response pathway that are not turned on as highly in smokers who get lung cancer. And Doug published a paper where he showed a polymorphism in the promoter region of one of those genes called aldo-keto reductase 1C1. So we have begun to look at what's upstream of the gene expression changes but not done a whole lot in that space yet.

DR. PRASAD: And, Dr. Bell, just one question. It's very interesting that you showed differential levels of AHRR methylation between T cells and monocytes. One is 4; another is 25, if I recall correctly. Do they metabolically -- if you look at the cells separately, are they metabolically equally

competent? Is there a reason why AHRR gene methylation levels are different? Can you correlate to the metabolic activity or some other factor, perhaps?

DR. BELL: Yes. We really don't know why there are differences we have between the cell types. We think one hypothesis is that the myeloid lineage has a rapid turnover and that the effect is basically related to a response to a proliferative -- a proliferative response.

And so typically myeloid cells are rapidly turning over hours to days, whereas the lymphoid cells are much more longer lived and are only proliferating less. There's, you know, a big response, immune response going on. So that's sort of our concept. We're really in a process of doing much finer separation of cells to identify which are the cell types that are changing. For example, we do see some changes in T cells. As you said, it's small, and we believe that small difference that we observe is due to a small subset of T cells that change quite a bit. So in other words, say 5% of the T cells are the ones that are the sensitive ones that are changing a lot.

DR. PRASAD: May I have a final question to Dr. Miller? Dr. Miller, you talked about pluripotent cells in your presentation. You went through too fast. Can you speak a

little bit about them, please?

DR. MILLER: Could you repeat the last part of the question?

DR. PRASAD: Can you just talk about those pluripotent stem cells and how they respond to smoking?

DR. MILLER: Okay.

DR. PRASAD: And how do they correlate to EPCs? Did you look at any correlation in blood?

DR. MILLER: No. You know, we're just beginning this work. We found that, for instance -- you know, there's been some discussion of biopsy versus brushing. Biopsy gives you just probably 10 times more of these progenitor cells. So I think if you're going to study them, you probably do have to get more invasive, which of course a bronchial biopsy is very invasive. We've found, you know, that they do seem to -- some of them, you know, as I showed, respond to iloprost by, you know, becoming multipotential. It's interesting that in the iloprost trial, the current smokers didn't respond at all, but it doesn't seem -- once you get these cells out of the smoker and into a dish, they do respond irrespective of the smoking status. So whatever that suppression is may be a very -- you know, sort of a very short-lived suppression.

But I think it's -- yeah, it's got a lot of potential. Some, but not all, of the mutations that we see in the dysplasias in the patients that we have studied are carried over into the cultures. But, you know, for instance, we have one patient who had a p53 mutation that we saw in the -- in his airway but not in his cultures. So, you know, it's not a perfect system by any means.

DR. DRESLER: Sir.

MR. SENGUPTA: Tapas Sengupta, RAI Services Company.

My question is kind of a two-part question. The first one is have you looked at the total methylation levels? Do they change? And the second question is have you looked into the housekeeping DNA methylizers that continuously work on keeping the methylation status? Have you looked at any change in them?

DR. BELL: I can answer the second question very quickly. No, we haven't looked at any of the methyltransferases or the TET enzymes, which demethylate.

Total methylation is a -- I mean, it's sort of where methylation started, in terms of where people started to look at the methylation of DNA, and we have not looked at that. We have developed some methods to kind of get a surrogate level about total global methylation by sampling, by basically

analyzing the results from the arrays and the sequencing.

But at this point, well, we haven't really done it with our new data yet, and that is a -- the global methylation levels are essentially dominated by the methylation levels of the repeat sequences in the genome.

And so in case-control studies and other studies where they observe a difference in global methylation, this is sort of driven by this sort of change in the repeat sequences in the genome, and it's a really interesting observation which no one really understands yet. And we think that as we get more into this whole genome sequencing and understanding the relationship between methylation and the phenotype or transcriptome, we'll begin to understand what that global methylation measurement means. Yeah.

MR. SENGUPTA: Can I ask another? Another thing is I was thinking about the methyl donor, the S-adenosylmethionine. You know, it's a very critical molecule, and as you know, its synthesis level is kind of related to age and disease and all of that stuff. Has there been any work looking into that?

DR. BELL: Well, interesting. The first study that we started with was to look at folate levels in relationship to methylation. And so folate is really the origin of the methyl

groups, one of the methyl donors for methylation at all levels. Well, in any case, we weren't really able to detect differences, that differences in folate intake, dietary folate, made different -- observed differences in the methylation levels at a global level.

But for various reasons, it's difficult to do these folate studies now because everyone does folate supplementation in various forms. And so probably many people -- most people probably get too much folate these days, I think. I'm not really sure about that. But so, no, we haven't really been able to measure a connection between folate levels and methylation.

DR. DRESLER: I'm going to go back to a question that I had started, Steve, with you and asking about the Japanese and they smoke different cigarettes. But again, the cigarettes are quite broad. Not all cigarettes are created equally, right? So they're different.

Has anyone started to look at any of the variability that's going across in these very fine-tuned, complex assessments that you're doing, and being a little bit less gross about the delivery of the smoke? Is anybody starting to look -- you know, we look at how many cigarettes per day,

perhaps not as granular as perhaps we should be. Or the duration. But how about what type of cigarette is doing what?

DR. SPIRA: I mean, it's a great question. In our prior work on the airway transcriptome, we hadn't even collected the type of cigarette. We collect everything else but that. I think a few years ago we recognized what you were saying to be a very important area to look at. So all the prospective studies, including the one we're doing with the Department of Defense, we are taking very, very detailed tobacco-smoking histories, including brands of cigarettes. People tend to be loyal, so that's the good news.

Where we're having a particular challenge is in our electronic cigarette study, trying to get ahold of what exactly they're using in terms of an e-cigarette, and we're finding now, you know, third-generation products essentially allow the user to mix and match ingredients. And so it's very difficult for us to get an accurate representation of what's actually being inhaled in that vapor.

So it's an important question for cigarettes and actually maybe even more important for electronic cigarettes, trying to get better resolution as to what's being smoked.

DR. DRESLER: Steve.

DR. HECHT: In the Multi-Ethnic Cohort Study that our collaborators in Hawaii and USC carried out and that we worked with on some of the data that I showed you, they didn't have really data on different types of cigarettes, but we have a huge amount of biomarker data. But your question is very important, and Dorothy Hatsukami and I have carried out many different studies on different products and different types of cigarettes, including one that we're doing now on electronic cigarettes. So it's an important point.

DR. SARKAR: Hi, my name is Mohamadi Sarkar, from Altria.

So I have a general question for the panelists. So we heard some really cutting-edge science and, you know, as you said, Dr. Dresler, drilling down to the molecular level. But if we kind of just look at it at a 30,000-foot level, what does this mean and would you, as the panelist, be comfortable making some conclusions or inferences from these changes, whether these are biomarkers for cancer? So if you see significant changes in the expressions, does that mean that you would get a lower risk for lung cancer?

DR. SPIRA: That's a tough question. I guess I'll be the first to take a stab at it here. At least in our work, what we think -- and we don't have proof of this, but we think the

field that we're studying is a measure of risk, not actually a measure of the presence of malignancy. The reason we think that -- and we're not certain of it, but the reason we believe that right now is it appears, regardless of where the tumor is, your right or left lung or whether it's central or peripheral, when we sample the right main stem bronchus for the biomarker, the performance is the same; that is to say, it sensitively picks up tumors regardless of location or proximity to where we're brushing, arguing that it's unlikely some local factor being produced by the tumor altering the local field but rather either something systemic or a predisposition.

And, in fact, the predisposition piece, given that the specificity isn't very good, we argue that our marker may not be measuring whether that nodule today is cancer but measuring whether that individual is at high risk. And so even though that nodule turns out to be benign, 5 years from now this may be someone who develops lung cancer, and that's what our marker is actually measuring.

Now, the only way to sort of prove that out is to do studies where -- and we're hoping to do this soon, where we actually take the tumor out of a patient and resample the field and see how it changes over time with treatment. And we

haven't done those studies.

So again, your question is a good one in terms of our biomarker. We believe it's measuring risk as sort of a gene by environment interaction. But as I said, the studies to prove that out are still pending.

DR. DRESLER: So are any of the markers that we've heard about this morning, are any of them relevant to smokeless tobacco and oral cancer?

DR. HECHT: I would say NNN is. I mean, that's what all the data would indicate. I mentioned the study from the Shanghai cohort because that was esophageal cancer. But I mean, we know that NNN causes oral cancer in rats. It's readily measured in urine. And I think with respect to oral cancer, NNN will be an important biomarker. But, you know, the studies are harder to do because the disease is less prevalent.

And, you know, the other thing is the DNA adduct studies. I think the oral cells are a great source of DNA for adduct studies, and I think this can go forward now. It's practical.

DR. DRESLER: Sir.

DR. CLARK: Bruce Clark from Philip Morris International.

Just a point of clarification, I think, on the question that Dr. Hatsukami asked you, Dr. Hecht, earlier about the use

of exposure, biomarkers of potential exposure biomarkers, and your comment related to nicotine equivalents being -- gives you everything. From an industry point of view, when we're looking at modified risk tobacco products, which may not have the same and do not have the same relationship between nicotine exposure and other HPHC exposure, are there others that come to mind that are useful as exposure biomarkers that we could look at without having the potential harm?

DR. HECHT: Sure. NNAL, phenanthrene tetraol. I mean, NNAL is going to tell you about NNK. NNN will tell you about NNN. Phenanthrene tetraol will tell you about polycyclic aromatic hydrocarbons. So I mean, if you're looking at modified risk tobacco products or let's say you're comparing cigarettes with smokeless tobacco, so you know, the volatile organic compounds, all the mercapturic acids, acrolein, there are going to be huge differences between these different products.

So you're right. I mean, when you're comparing different products, it's not so simple as nicotine equivalents. I mean, if you're comparing smokeless tobacco --

DR. CLARK: Yeah.

DR. HECHT: -- to cigarette smoke, you're going to get

similar results for total nicotine equivalents, and we know the risk is hugely different.

DR. CLARK: Right.

DR. HECHT: So yeah, you have to take into account the whole thing. But I still think total nicotine equivalents, used in the right setting, will be a biomarker of harm.

DR. CLARK: And the measures are often standardized against nicotine exposure, so you have a relative comparator as well. Thank you.

DR. DRESLER: So one last question that I think I know the answer to. So we asked about smokeless tobacco. How about waterpipes? So people are looking at e-cigarettes. Who's looking at this for waterpipes?

DR. HECHT: So we've looked at waterpipes. There's a person at San Diego State University, named Nada Kassem, who is really into waterpipes, and she's done several studies. We've done the biomarker work on that, and the one thing that stands out -- and the Benowitz group observed this also -- is high exposure to benzene in waterpipe users relative to other compounds. So that's a significant thing to worry about. The other thing that comes out of it is secondhand smoke exposure in the family when waterpipes are used at home.

DR. SPIRA: We haven't done the studies yet. We have actually a former graduate student who worked with Dorothy and Steve who's now back in Pakistan, and she's got a cohort there that we're trying to connect and actually get at least nasal samples, probably not bronchoscopy.

The other exposure we're really interested in looking at is marijuana and how that impacts airway gene expression, and we're trying to put together that kind of study. So I agree, there are additional approaches that can be brought to bear.

DR. DRESLER: Very good. Any other last questions? Because I think we've about caught up, so it is lunch. It's still a little early. So it is a quarter to 12:00. So let's come back at a quarter to 1:00, and we'll start the next session.

Thank you very much, our speakers and panelists, for this. It was an interesting session. And we'll see you all after lunch.

(Whereupon, at 11:47 a.m., a lunch recess was taken.)

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

(12:47 p.m.)

DR. DRESLER: Shall we get started? I think most of us -- it looks like most of us are inside.

So we are coming up on the afternoon for Session 5: New Areas of Research. And the first presenter will be Dr. Gayle DeBord from the CDC National Institute of Occupational Safety and Health, speaking on Use of Biomarkers of Early Effect in Occupational Exposure Limit Setting.

DR. DeBORD: So I'd like to thank the organizers of this meeting for inviting me to present today. And I have no conflict of interest.

And so this morning we really learned a lot about specific biomarkers, and I'm going to be more at the 50,000-foot level and really to see how we can use these biomarkers, how we can apply these biomarkers.

So an occupational exposure limit is the acceptable concentration of a hazardous substance in air, in workplace air. So really it's a critical component of risk management. There are many organizations that set occupational exposure limits, but they really don't provide a lot of information on how they do that and what their rationale is for having a

certain level. So, you know, they don't really talk about what the underlying risk -- the quantitative risk assessment is that they employed.

So what the National Institute for Occupational Safety and Health did was put together a special journal article in the *Journal of Occupational and Environmental Hygiene*, a special section, and there are 10 articles that pretty much describe the state of science for occupational exposure limit setting. And some of the focus is on early biomarkers of effect, epigenetic changes as well as what's the decision logic as far as exposure assessment is concerned.

So really, you know, the basis of occupational exposure limit is risk assessment. And so this is the four-step process for risk assessment. Once you know what the hazard is, you have two components that are very important. The first is exposure assessment, which is what are the exposures and how are those exposures incurred, and what are the conditions that are underlying those exposures?

Probably what's more important for this particular workshop, when we're talking about biomarkers of potential harm, is what is the dose-response assessment? You know, so what are the health problems? What is the dose response

that -- you know, what does the curve look like? And those two, the dose-response and exposure assessment, are combined really to look at risk characterization. You know, what's the magnitude of the risk from those exposures?

And it may be that the risk is decreased. So there are some benefits. Such as with the antioxidant exposure, you might actually have some benefits to that. Or what is the -- you know, or it also can have an increased risk.

But really, you know, as far as biomarkers of early effect or biomarkers of potential harm, what we're really going to focus on or where they're going to be most impactful is probably in the dose-response assessment.

So just some important terms of risk assessment is mode of action and mechanism of action are sometimes used synonymously in the literature, but they're very different with risk assessment. Mode of action is a higher-level organization, where mechanism of action is at the molecular level. And, you know, just kind of as an example, if you look at pathways to heart failure, there are basically four pathways you can go down that might be considered the mode of action, one of those pathways. Or even if you look at insulin resistance, that might be a mode of action where the underlying mechanism of

action may be, you know, the receptor sensitivity.

So as we're looking at biomarkers of potential harm, you know, in risk assessment we have different levels of exposure. We have the no exposure -- no observed biological effect level. We also have the no observed adverse effect level. And so I think this is where the biomarkers of potential harm have the greatest potential use, especially in that NOAEL level. You're looking at an adverse effect, but there still may be effects that have no clinical relevance, which is at the NOEL level.

So I guess it goes back really to what Dr. Bell was talking about this morning, when he was talking about hypermethylation. Is hypermethylation, is that an adverse effect? Is it a frank effect or is a low-level effect, or does it have no relevance whatsoever? It's probably going to be more in some kind of adverse effect.

So how do we fill in the blanks to improve occupational exposure limit? Well, one is we need more studies to look at comprehensive disease assessments. You know, what are the biological pathways that lead to disease, and are there biomarkers that we can use? The exposome is a concept that Chris Wild came up with about 10 years ago, and I think this is really going to make -- it makes life more interesting because

if you look at some biomarkers of effect or -- you know, that's really going to be in our unique characteristics.

You know, we have -- people are exposed to all kinds of things throughout their lifetime, and so that's really going to make the identification of certain biomarkers of potential harm more difficult because we all come -- you know, humans are very heterogeneous. And so that really kind of makes things a lot more muddy as far as trying to identify, you know, specific biomarkers.

But along with that -- and I think we talked about this earlier today, is that there are biomarkers of exposure that can act as potential biomarkers of harm, and so we do need better comprehensive exposure assessments.

So legacy biomarkers. These are biomarkers -- and I consider this targeted. You know, we know what we're looking for, and so it may be something like DNA mutations, or it may be an antibody. You know, the DNA mutations could also serve as potential biomarkers of harm. We also have response biomarkers. And Dr. Shiels this morning talked about a whole battery of biomarkers that she looked at for inflammation, and so that's a very targeted look.

But the omics technologies are some biomarkers that I

think will actually -- you know, well, Dr. Spira had talked about one in particular that I thought was very interesting. The omics biomarkers are going to be more discovery at this point. But like I say, Dr. Spira turned it into something that was more specific and more targeted.

But we want something, we want biomarkers. As we're trying to discover new biomarkers for exposures and for effects, you know, we want something that is going to give us a lot of information with a small sample size that we can do quickly and, you know, fairly inexpensively. And so the omics technologies kind of, you know, hit that particular mark.

So the biomarkers, you know, we've talked about some of those this morning, the adductomics, you know, the DNA adducts, the transcriptomics, proteomics. Those are all -- you know, those are just some definitions. But there have been some proof-of-concept studies using toxicogenomic data and trying to establish, you know, a benchmark dose or a dose-response curve using toxicogenomic data, and those have been actually fairly successful. So we really think that, you know, as far as risk assessment is concerned, the use of these omics has a lot of potential, a lot of promise.

So how can biomarkers aid in the prevention of

occupational disease? And also how can it prevent other diseases, not just occupational diseases? You know, one of the things that we can do with biomarkers is that we can look -- see how we can look at the different individual factors. You know, what are the factors that interact to lead to an occupational disease? What are the biological pathways that -- you know, that disease progresses?

And so, you know, it can be individual factors such as genetics, or it can be what your environment is. And your environmental exposures are not just specific exposures to you; it's also general exposures. You know, what's your access to healthcare? What's your income level? Where do you live? Because sometimes where you live is -- you know, gives you a whole different set of exposures.

And then the types and timing of exposure, which I thought was really interesting, Dr. Spira, what you had this morning when you were talking about the three-cigarette bolus. So is it that one bolus of exposure, or is it the chronic low-level exposures that someone might have over a lifetime? So is it someone smoking, you know, 5 cigarettes, 10 cigarettes a day across 40 years or -- you know, and according to what you may have indicated or suggest, that maybe it's not. Maybe it's

just this short bolus of three cigarettes within a short period of time.

You know, we run into the same problem in occupational health, where we have exposures where -- you know, in manufacturing facilities you have these low-level exposures that occur across 30, 40 years.

But then we also have these peak exposures in different industries such as with dry cleaning, when they open up the dry cleaning machine and they get a bolus of whatever the dry cleaning agent is. It used to be perchloroethylene. Now they've substituted that. Or even oil and gas extraction work where they have to check the tanks, and so to see what's in the tanks. So they open up the thief hatch, and they get this bolus of volatile organic compounds coming out. So are those, you know, 30-second, 1-minute exposures important, and if so, how are they important? These are things that we just don't know. We don't have a lot of data yet on those kinds of things.

But also the timing of exposures. You know, if you're exposed when you're younger, you know, when your metabolic enzymes aren't mature, your immune system isn't mature, does that make you more susceptible?

And certainly we have seen that there are certain chemicals, for example, diethylstilbestrol. You know, mothers took that back when they were pregnant, and now if you look at their children, you know, they now have reproductive deformities as well as reproductive cancers. So the timing of exposures, you know, when those exposures occur, can also change.

And then also we have these roles of mixed and cumulative exposures. You know, certainly cigarette smoke is a very mixed exposure. And so trying to find a biomarker of potential harm, you know, I don't think you're going to find one. I think it's going to be a battery of biomarkers. But if we can use biomarkers, we can improve risk assessment, because as we get more into the biological pathways, we can decrease the uncertainty that we see.

So maybe not so much in smoking, but for the rest of us in the world, we do a lot of extrapolation from animal studies to humans because there isn't a lot of human data. I think there is more -- I think the smoking arena is more mature than a lot of the other chemicals or other hazards, you know, that we face. And so we do a lot of extrapolations from animal studies to humans.

And we have to really look carefully at that because there are a lot of challenges that come with that extrapolation. You know, we have species differences, you know, the difference between a rat and a human. If you look at arylamine exposures, you know, the target for most arylamines is bladder cancer in humans. But for the rat, it's liver cancer. So, you know, there's that kind of difference. And then if you look at animals, they're very homogeneous and -- you know, as compared to the heterogeneity of human populations.

Animal models are usually short term as compared to this really complex lifetime exposures that humans experience. There are a lot of uncertainties as far as gaps. We have to worry about the different types of routes of exposure. In occupation, it's mostly inhalation, which is usually how animals are exposed. But there's also ingestion, there's dermal exposures. And so there's a difference in exposures between animal studies and human studies.

And then we also have to normalize that data because we have differences in body size, metabolic rate, and toxicokinetics.

So there are a lot of challenges when looking -- plus, animals are exposed at very high levels, and in humans, we're

really interested in that lower part of that dose-response curve.

But there is great potential for these early biomarkers of effect and response, and one is mode of action, because we can look at these biomarkers and we can try to figure out what pathways they're going down. So that way we can try to, you know, develop or identify biomarkers that might go down that pathway.

We can characterize inter-individual variability. We can reduce the uncertainty from between animal data to humans because we're now going to be looking more at human biomarkers. We can test our hypotheses at more relevant doses. And then we can look at these lower levels of dose, you know, down at the dose-response curve, that are more meaningful.

But there are challenges in -- you know, with these. Computational toxicology and systems biology -- and systems biology is basically just the biological pathways, you know, that a chemical might go through that might affect a particular disease process. So these are still evolving. We really don't know exactly how things work just yet, you know, down at the molecular level.

We have very few biomarkers that have been validated in

human populations to any great extent, and that limits their use, really. You know, you can validate a biomarker in the field, and you certainly can validate it for a clinical response, but you really have to look at that human population. You have to test it in the population that you want to use it on. And also what is the appropriate interpretation of the biomarker?

There are a lot of databases out there, there are a lot of biomarkers out there, there is a lot of information. But how to do you mine that data to get relevant information that's going to be important to be used in risk assessment and for setting occupational exposure limits?

So the future. I think the future is pretty bright and -- you know, because what we can do is we can identify sensitive, relevant endpoints that can be targeted, and that's our biomarkers of potential harm.

You know, they may be preclinical, and what may happen is that, you know, we go back and say, okay, these biomarkers are biomarkers of inflammation. We don't really know what the process of inflammation to disease is yet, but that's okay. That may be okay.

You know, until we can fill out the whole exposure-to-

disease continuum, you know, I think that we're looking at identifying preclinical biomarkers that may or may not have material impairment. But they still are markers that something is amiss, and that may be biomarkers of exposure that -- you know, some biomarkers of exposure may be able to do that. But then also these biomarkers of early effect, like I say, may be preclinical and will be useful in helping us to use in our risk assessments to actually decrease our uncertainty of what some of our risk assessments are.

You know, in risk assessment there's a lot of conservatism. And so if we can be more accurate in what the exposure is and what the effects are, we don't have to use so many uncertainty factors. But the basis of it all, though, is to try to prevent disease, and so that's how we are -- that's where we are.

And so I'd like acknowledge my co-authors on -- we actually wrote the fifth paper in that series. There were 10 papers. And Steve Edwards is in the audience too.

Thank you all. And we'll take questions at the panel discussion. Thank you.

(Applause.)

DR. DRESLER: So our next speaker was one of those

co-authors, I believe, if I saw that correctly. So Stephen Edwards, Dr. Stephen Edwards from the U.S. EPA will be speaking on Biomarkers and AOPs: Bridging Toxicity Pathways and Regulatory Decision Making.

DR. EDWARDS: So thanks, everyone, for sticking it out until the last session. So I have no conflicts of interest to declare and have the standard disclaimer that everything I say is my opinion and does not reflect EPA policy.

So what I would like to go over today is I'm going to mainly be talking about adverse outcome pathways, but I want to start by kind of putting that in context. And a lot of this will follow on nicely from the talk you just heard. And then I'm going to kind of break into two pieces.

So one is going to be on the AOP, and I'll spend probably most of my time there, really. And the idea there is thinking about how to structure the information, like all of the stuff that we've heard about for the past 2 days, and can we use this AOP construct as an organizing tool? And then at the end I'll come back and more explicitly talk about the biomarker data. And I've made some notes as the meeting has progressed, so I'll try and make some points as to how it might tie back to some of the previous work you've heard.

But just to put this in the context of what we do at the EPA, so we kind of wind up with this gap between the things that we can regulate, which are chemicals in the environment, and the criteria that we use for regulation, which tends to be -- in our case we have to deal with humans and wildlife. So we deal with individual populations and sometimes even communities when we're talking about ecological risk assessments. And so there have been several frameworks that have come out to kind of facilitate that. The thing that prompted a lot of the work I'll talk about today was an NRC report back in 2007 where they came up with this idea of the toxicity pathway, which was a cellular response that could be monitored in vitro that would be indicative of an eventual adverse outcome.

But the trick with the toxicity pathway is you have to figure out what that adverse outcome would be and what the relationship is before you can use those in and of themselves.

The mode of action framework, which has been around for quite a while and you heard about it in Dr. DeBord's talk, is a great bridge for that. But the one issue with the mode of action for interpreting toxicity pathways is mode of actions, as they had been done, were typically chemical specific. So

you accumulated all the information you had about a chemical, you developed your mode of action assessment, and then you moved to the next chemical.

And so if we're looking to do in vitro screening and relate it to adverse outcomes, then we really need something that does not -- where we do not have to go through that whole process every time.

So in 2010, a group of my colleagues up in Duluth, Minnesota, who work at EPA, published a paper on an adverse outcome pathway where we basically focus on the pharmacodynamic parts of the process, but leaving out more of the chemical-specific information. And the advantage to this is now we can tie these molecular-initiating events to these toxicity pathways. And so now we can assay a lot of chemicals on high throughput and then try and use this adverse outcome pathway to interpret what the downstream adverse outcomes might be.

More recently we've defined an exposure counterpart to that, and so I'm not going to go into that in great detail today, but I just wanted to mention it so that people would know that we are thinking about the exposure side of things. So just briefly to cover that one.

The AEP is built off of the constructs that we pulled

together with the AOP. So as you'll see when I get to the AOP part, we break things into key events, which are nodes, and key event relationships, which are edges. And consistent with the way key events are treated in the adverse outcome pathway, the nodes really represent the measurable components, and the edges represent the changes that take you from one node to the next. With the AEP we are focused on chemicals. That's the whole point. And we end with this thing called a target site exposure, which is actually intended to match up with the molecular-initiating event, which is how we connect to the AOP.

As you might imagine, when we're thinking about exposure, the AEP does not explicitly include toxicity. Just because you're exposed does not mean it's toxic.

And so if you think about this in terms of the mode of action that you just heard about, the mode of action really breaks down into a smaller piece of this puzzle where you eliminate the external exposures and you really start from, you know, either in a lab animal study, the dose that was given, or in a field study, you know, the exposure at the human receptor.

And so then you deal with the pharmacokinetic issues that get you to the target site exposure. And then you start with the macromolecular interactions, so the interaction of that

chemical with the biological system, which then leads through a series of events to outcomes that can be used for regulatory purposes.

So the AOP has very similar components. Key events are nodes or a change in biological state. They must be measurable, and they must be essential for the progression down the pathway.

We have two special key events. There's our molecular-initiating event, which is the initial point of chemical contact, and an adverse outcome, which would be really at one of these two levels, which is basically anything that we would consider adverse for human health for a human risk assessment.

And the relationships are where we capture the evidence for supporting the AOP. So between each of our key events we're going to capture the evidence for that pair. And the reason for this -- and this another place where the AOPs diverge slightly from mode of action in that we do -- in a mode of action, you kind of discuss everything together, whereas with the AOP we get very modular about how we do this.

And again, the idea here is reusability. So what we want to get away from is having to do everything from scratch each time we have a decision to make. So if we make these decisions

at each key event pair, then we actually do another assessment. Even if it's not using the same AOP, if it's using parts of that AOP, we can use the knowledge that has already been gained to help us with the new risk assessment. And as I mentioned earlier, the AOP does not explicitly include chemicals. So it's entirely focused on the biological processes.

And then, to complete the continuum, we can connect the AEP to the AOP, with the AEP dealing with how do chemicals get to a target site, and the AOP dealing with is a perturbation at that target site sufficient for an adverse outcome?

And so the first criticism of this is okay, you've got this great construct. But now how do we -- you know, it's going to take us forever to get all of our AOPs. So what do we use in the meantime? And so what we've done is we've actually gone through all three areas and worked out methods for addressing things in the short term, and then more data-intensive methods for addressing things where the precision is higher.

I'm not going to talk about the ADME or the exposure today. I have colleagues at the EPA that work in those areas. But I'm just going to talk through the AOP as an example of how this process works. And the AOP work really is done under an

OECD program, which was started back in 2012, and it's run by the Extended Advisory Group for Molecular Screening and Toxicogenomics. And this group offers guidance and training for creating AOPs as well as an international knowledge base to capture the information from the AOP authors.

The knowledge base itself I'm not going to spend much time on, but it consists of a wiki, which collects the information regarding the weight of evidence; an AOP-Xplorer, which allows you to peruse AOPs in a network context, and I'll come back to that more in a moment; an Effectopedia, which captures the more quantitative aspects of the AOPs. And then of course we have a hub, which will be out soon, which connects the individual modules together.

So there are five principles to AOP development. The first one, AOPs are not chemical specific, I've already covered. The second two I covered in the introductory slides with KERs, key events and key event relationships. And actually three and four lead into a concept which is central to the next few slides that I will present, and that is that we understand that a linear AOP is not an accurate description of biology. Biology is much more complex than that.

And so one of the criticisms we get often is, well, you've

oversimplified things, and it's never going to be useful. But the idea is we simplify things so that we can describe it, and then we merge things back together for the actual use. And I'll talk in great detail about, you know, things that we are doing to make sure that these two points are made. And then AOPs are living documents, and that's rather central to everything that I'm going to talk about.

And so we've kind of loosely defined three milestones along the way and gave them labels just so we could talk about these things. But really, this is a continuum. And we actually hope -- and if it were a different talk, I would actually try to convince you that AOPs can actually be useful all the way along this continuum. So we do not have to get an AOP all the way out to where we can make quantitative predictions before we can use that AOP in a risk assessment context.

So putative AOPs we pretty much define as we put together these boxes and arrows and we have a general concept for what the AOP is.

Formal AOPs have undergone a formal evaluation based on the guidelines set forth by the OECD development program. At the end of this process, you can actually submit these to the

OECD for a formal review. And if they are approved, then they will get an OECD endorsement and will be published by the OECD.

Quantitative AOPs take this even a step further and actually describe the -- actually the response-response characteristics, because we don't care about the individual chemicals. But you look at the dose-response concordance across the different key events and work out, if key event 1 changes by X percent, you know, what would be the quantitative change expected in the next key event downstream?

The knowledge base that I mentioned actually is being developed to support all of these stages of development. And so the AOP Xplorer helps you navigate the early stage putative AOPs and eventually will provide even more information and helps you to define the AOPs or at least a hypothesis of what the AOPs would be like.

The AOP Wiki is more designed around structured forms for doing the evaluation. So it is tied very closely to the OECD handbook for AOP development. And Effectopedia, as I mentioned earlier, is more focused on the quantitative aspects.

So going back to Principle 4, which is our AOP networks, the first thing that we've done to facilitate AOP network generation is the knowledge base itself actually stores these

things as networks. So as you are entering the AOPs, even though you may be describing them as separate entities, if you enter them according to the guidelines, they will actually get combined, if it makes sense.

And this just shows one example looking at estrogen receptor signaling. So we could have an estrogen receptor antagonist, which is described as AOP 30 in the wiki. We could also have a chemical that inhibits the aromatase enzyme. So this decreases the synthesis of estrogen and therefore circulating estrogen levels, which eventually feeds into the estrogen signaling. That's a separate AOP. And we can also have androgen receptor agonism, which as a side effect decreases the amount of estrogen, circulating estrogen. So these would be three and actually are three separate AOPs in the wiki.

But because they share multiple key events and key event relationships, the way that information actually gets stored in the wiki is this. So the AR agonism has a couple of specific key events and then feeds in at this point with the aromatase because this is at the point that estrogen is being synthesized. And then that actually feeds into the estrogen receptor pathway, starting at the point of changes in estrogen

receptor signaling.

So if we continued down this pathway, eventually we could just keep describing individual AOPs, and eventually we would get to something that describes the overall system just by assembling these things. However, that's likely to take a much longer period of time than we want to wait.

Oh, before I move on to that part, just to give you the current status of AOP development. So we've got about 100 putative AOPs in the wiki at the moment, and most of these come from a project we have internal to EPA where we're literally getting things to this stage and stopping. And so the hope is that we can get crowd sourcing and have people from around the world help us in developing these further. So if anyone is interested, take a look.

There are 18 formal AOPs that are currently in the process of OECD review. There are about five, I think, that are up for endorsement now and should be out soon.

And then the quantitative AOP development is a much bigger endeavor. It takes quite a bit of data and quite a bit of time, so there are not that many of those currently. But there are still many assays that do not have AOPs. So the ToxCast program that we deal with has over 400 targets at the moment,

and we only have AOPs for about 40.

We also understand that we don't have all the targets we need, so we're looking for new AOPs. And so one of the projects that my group is involved with is can we do things in the short term to speed up this process?

And so basically, in short, I'm not going into this today in the interest of time, but basically what we're trying to do is reverse the process that I just described. So now we're going to use data mining methods to infer what this network might be. We're then going to use network traversal algorithms to extract out these individual AOP networks. And then these can actually feed into, which I think is the next slide, the expert-driven process that I described to begin with.

And so the hope is that by inferring these networks and doing this, we can both speed this process -- because now, instead of giving them a blank slate to try and pool AOPs together, we're giving them something to respond to. So we think it should enable the expert review to move more quickly, but it also gives us other things for some of the high-throughput approaches that we're taking. It gives us something in the meantime that we can use for decision making.

And obviously, as things move through this process, we're

going to get feedback on which of these things were correct and which were incorrect. We can feed that information back into our data mining approaches and hopefully improve the ability to predict this network.

And so again, the point here is that we're working on ways to get a lot of AOPs with very limited information. We're also working on ways to systematically move those AOPs down the process. So the idea is we can actually prioritize these based on the decisions, the needs, information needs for a given decision. And so AOPs that aren't sufficient at these higher levels would then become a priority for moving into taking the effort to move them into the lower levels.

And like I said, we have parallel efforts both looking at ADME -- my colleague who was here at the last workshop, Cecilia Tan, is working on the ADME. We have other colleagues who are focused more on the exposure side.

So then how do biomarkers fit into all of this? So as I mentioned at the beginning, the AOP really is intended as an organizational framework. And so what I would like to see is, all of the information that you have presented over the last 2 days, I could easily see how we could fit that into an AOP framework.

And so whereas we tend to focus on these high-throughput assays and things that we can run in high throughput, the same things apply for these early stage biomarkers that Dr. DeBord talked about and that all of you have been presenting over the past few days. And then, obviously, we did have the talk yesterday on the in vitro approaches, which would fit into this framework.

And so the idea is that you can build up these AOPs and take advantage of the existing knowledge. And then for each of these key events, you can tie your biomarkers to those individual key events.

And so now you have documented evidence for, okay, if I have perturbation of this key event, say, somewhere in the middle, now I have documented evidence for what that likely means in terms of an organism or a population response. And I can just keep backing this up until I get to something that can either be measured in high throughput, which is our goal -- one of the things that came up yesterday was this idea of the temporal nature and how a lot of these things, it takes a long time to manifest.

Well, if we can take all of the data you have and establish causal connections between some of these earlier

responses and those outcomes that do take the time, then we can actually start using some of these possibly for the newer evaluations of the alternatives.

And so that's really where I think this fits into the biomarkers is as you assign your biomarkers to these key events along the way, you take advantage of the fact that you have a wealth of information from decades of smoking research that you could put into informing the weight of evidence connecting these key events, and then you could use this for the decision making.

Now, what it doesn't address, which came up yesterday, are the unknown unknowns. So if you've got an alternative and it all of sudden is doing something that cigarette smoking did not do, we won't address that. But if it's simply a matter of how things perturb those existing pathways, you might be able to leverage a lot of data that you already have to save you from having to collect that data again.

And so one of the things, our biomarker work kind of has taken a backseat to really getting the AOP work off the ground, but we do have a concerted effort within the EPA to look into biomarkers. And we really are focused on the target organism, which we say since we have ecological species. In the case of

this meeting, you know, the target organism would be humans.

And so we really do have a focus there, and we're hoping to get biomarkers that are measurable in an accessible matrix, like blood or urine, and then could be used for a variety of studies and could be used to complement both the lab animal data that we're collecting as well as the in vitro work.

And then, as I mentioned, we do have biomarker work. I'm not going to go into any of this in the interest of time, but I did want to alert folks. In my group, like I said, we tend to do inference based on large datasets. So we have a project where we looked at the NHANES data and found chemical associations with the different effects biomarkers. And the unique thing about this, over a lot of the data mining that's done in NHANES, is we did it comprehensively. So we don't give you an individual p-value for whether this is significant or not because we really don't believe those when you're doing a data mining exercise. But what we do is we give you relative probabilities because we looked at everything.

So we can say for chemical X, you know, the most associated biomarker was Y. And then it would be up to you to go and follow that up and determine is this a real association or is it just something that happened to fall out of the NHANES

data?

We've also done some work looking at effect modifiers. So we can take these AOP networks, and just as we can look for chemical mixtures and interactions among different chemicals, we can also look for effects modifiers. The nice thing about the AOP framework is it does allow us to connect in these effects modifiers at the level of the organization where they're most appropriate. So genetic susceptibility, you would model that at the molecular level, whereas some of the other things may come in at some of these higher levels of organization.

We had a paper a year or so ago on data-driven asthma endotypes, where we looked at gene expression in blood and tried to classify asthmatics. And we were actually -- it was quite successful in that we were able to recapitulate a lot of the current diagnostic criteria while also finding some interesting new -- potential new subtypes.

And so again, just to wrap up, the AOP framework is really designed to organize the information that you've heard about. And mode of action can really be thought of as a combination of AOP and chemical-specific information. And it really, I think, can address your conundrum where you had many years to get to

an outcome. It could enable more in vitro screening results, like you heard yesterday. And it does provide a framework for establishing biomarkers in a mechanistic context.

Just the quick acknowledgements. Cecilia Tan and Dan Villeneuve are colleagues who have worked on several of these areas. We have a whole project team on AOP discovery and development. Our KB has a work -- international working group; our biomarkers team as well as others. And even though I didn't talk specifically about any of their work, I have to acknowledge all of the folks in my group who do all the work while I'm out talking.

Oh. And if anyone is interested in the AOP concept, there is going to be a meeting, and they actually have a call out for questions to seed the meeting.

And I guess we'll do questions.

(Applause.)

DR. DRESLER: Our next speaker will be Dr. Christopher Proctor from British American Tobacco, speaking on Utility of Biomarkers of Potential Harm as End-Points for the Disease Relevant Assessment for -- you have the longest title. You win, you win.

(Laughter.)

DR. DRESLER: There it is. All right.

DR. PROCTOR: Good afternoon. First of all, I want to say thank you in particular to the Center for Tobacco Products for (a) organizing so well and (b) facilitating so well these workshops. They are something that really helps me in our research program. Every time I come to these, I learn a lot, and I bring that back into our research program. Our research program is trying to develop a next generation of tobacco nicotine products which are substantially less harmful than conventional cigarettes.

As a conflict of interest, I'm fully employed by British American Tobacco, which is a manufacturer of tobacco and nicotine products. And what I want to cover in this presentation is a little bit of an orientation. Fortunately, thanks to the last superb presentations, a lot of my orientation, particularly as I get further down into this deck on AOPs, has been given.

But I want to start with where we've come from, really, which is using the traditional biomarkers of potential harm to try and understand the modification that might be occurring when you introduce a reduced toxicant prototype cigarette, but move from there to where we've had to go, I think, to

supplement that data using AOPs and a lot of in vitro data. I'll use a few very fast case studies to do that. Then I want to just talk a little bit about bridging and how we practically apply some of this science to modified risk tobacco product applications.

I was really struck right at the beginning of this workshop by Dr. Hatsukami saying, well, we started this journey in 2001, so 15 years on, and the question is, so why don't we have a much better feeling for biomarkers that could operate for us in this field? And I think there's probably two answers to that. One is we probably haven't done enough on it, and secondly, that some of the things are really quite difficult to do. And we saw particularly, I think, in this workshop an awful lot of work on the disease end. So when people have the disease, then we're doing some incredible science to try and characterize that. And we saw a lot of that, particularly today.

But in the context for a modified risk tobacco product, we need to have healthy people in the kind of studies we're doing, and so the construct is a little different. It's really kind of markers which are much earlier along in the disease process.

Just to note. In terms of my definition there, I'm using

biomarkers of potential harm. But if you look at a lot of our literature, we talk about the kind of more neutral terminology of biomarkers of biological effect.

Okay, let's start with where we were sort of 4 or 5 years ago, trying to develop a cigarette that might be seen as less harmful than conventional cigarettes. And we did that by creating a product which had a whole bunch of technologies in it which reduced individual and groups of toxicants compared to cigarettes, and we wanted to test it in the way effectively the IOM has set out.

So, first of all, a short study of biomarkers of exposure, and then a 6-month clinical study of biomarkers of potential harm, and then trying to select those.

So we went into the literature, and eventually the qualifying selection criteria for them were -- well, the biomarkers would show differences between smokers and never smokers, and those biomarkers would be modifiable in the 6-month period that we were looking at in the study, and that they were somewhat related either to a specific disease endpoint or one of the pathways on the endpoint.

The products themselves look kind of like a cigarette, but they have a huge filter in them to try and get particularly the

vapor-phase constituents to be much lower than you'd get in any traditional cigarette. And they have modified tobacco in them as well.

If you look at the chemistry part there, you'll see that if you did a cumulative assessment of them, you would find that their toxicants were much lower than any traditional cigarette, though in some cases for certain classes of compound, we could only get them down a certain amount. If you look at the right-hand panel there, you see that at the end of the study, which was a 6-month study, as I said, we saw reductions in urine mutagenicity. But when we looked at the biomarkers of exposure, we found that, well, we have compliance issues.

So some people increased their consumption as they went through into the study, but we were able to find quite good distinction -- and here's a biomarker for acrolein -- between the red line, which is people continuing to smoke a conventional cigarette, to the blue line, which is the people being switched to the reduced toxicant prototype cigarette. And then we also measured both ex-smokers and never smokers to show that there's distinction, and interestingly really hard in the context of a cigarette to get your exposures down to where you'd expect or you'd hope an ex-smoker or a never smoker would

be.

But when we looked at the biomarkers of potential harm, we found really very little difference in the 6-month period between the test and the control. And that was true for quite a lot of biomarkers we've been talking about today, such as white blood cell count, sICAM, dehydrothromboxane B_2, and 8-epi-prostaglandin F2alpha. And the conclusion we came to at the end of this is it's really hard to modify a cigarette to find a reduction in risk, though I know we've heard in various presentations that the notion of actually even a small amount of tobacco smoke can cause effects, and the way we're modifying, people are still smoking 15 to 20 cigarettes a day, albeit quite modified in their toxicants profile.

It led us to a place where we thought, okay, we need to find a way in which we go much bigger in terms of the exposure reduction, and that is by going to electronic cigarettes and electronically tobacco-heating products where you can take the exposures down even further.

The other thing we decided was that, okay, biomarkers are going to be important, those clinical studies are going to be important, but they are really hard to do. To try and get people in a 6-month study, nonresidential, albeit occasionally

residential clinical studies, and get them to comply with what your ideal would be is just really, really difficult.

And so although I'm sure it's going to be part of the package you need to collect for assessing a new product, I think we felt we had to supplement it by other information.

The other thing in terms of our history, we're British, and both the British government and Brits tend not to like to do animal experimentation. So we've had a program for the last 10 years trying particularly to develop in vitro models that wouldn't require in vivo research. And so it was natural that we would go into that area.

Our research program has been a little bitty, so we were getting models with very little bits of a disease process which were being developed quite nicely, but they didn't seem to fit. So we eventually saw the light and thought, okay, adverse outcome pathway is the way in which you can frame a lot of this information, it really helps, and started to bring it together.

A lot of the focus that we have is obviously of defining the molecular-initiating event, then looking at trying to frame key events along these pathways, and then to explore, okay, cigarette smoke versus e-cigarette aerosol or other aerosols generated from next-generation products to see what you would

see in terms of a difference. The secret to the approach is -- and I think, as I say, it's complementary to more traditional approaches -- is that you can combine in vitro data, which you can collect much more rapidly, and you can collect in different conditions than you can into the clinic, with clinical data, and then put it within the framework of an adverse outcomes pathway.

So I'm just going to give you two case studies where we do that. In this first one we're looking at transcriptomics and metabolomics where the simple experiment is you expose the cells to smoke or to air, and you look at a variety of different expressions of that.

So you look at differential cytokine release, you look at the differential gene expression, and once you've done that in a series of different techniques, you just combine that together with -- what we use is IPA, ingenuity pathway analysis. It's to look at, okay, so where those differences are, how do they fit within the kind of pathways that already have been developed by looking at a variety of different things?

And you can see, in this case it's what you would expect, because we know what smoke does to you, that you're seeing lots

of signals going on in terms of tumor promotion, cell proliferation in the system. So you start with in vitro, and then you move to clinical samples. And in this case, these are blood serum samples from that reduced toxicant prototype clinical study where we had 60 smokers and 60 nonsmokers who we put in the front end, and we do very untargeted both mass spectrometry, liquid chromatography-mass spectrometry, and also NMR to collect data on the differences between those two samples, and then use PCA to work that up to see where you're seeing differences between the smoker groups and the nonsmoker groups. And that allows you then to map across, again, to the ingenuity (IPA) to see, okay, what things are really signaling as a difference.

If we move on to proteomics -- and here we do an awful lot of work with 3D lung cultures, that once they're in the laboratory, just as Dr. Tarran was talking about yesterday, you can culture nicely, and they grow and they live and they create this kind of air surface liquid similar to that of the sputum.

So in this set of experiments, what we're doing is letting these cultures grow, sampling the liquid and then seeing, using in this case LC-MS/MS, exactly what's happening when you get them exposed. And you can compare that then to healthy sputum

to see, okay, is what you're creating in the laboratory at least similar to what you might get in healthy sputum induced from volunteers? And what you find after doing the analysis, that you get a pretty strong overlap of the -- of common proteins between what you create in the laboratory and what you actually find in humans.

But then, of course, you want to switch over to clinical studies. In the clinical studies here, we've collected induced sputum samples from smokers and nonsmokers, analyzed those again, and you're seeking again for differences between the nonsmoker and smoker level. You can see that quite a few genes are quite different between the smoker and the nonsmoker.

And what you can do after that protein identification is then you can kind of quantify it up using multiple reaction monitoring. And what you find there is you get some candidates that you might then want to put into a longer-term clinical study. Actually, in this case we're seeing over a threefold change with the smoker sputum compared to the nonsmoker sputum.

Talking about the OECD program, ourselves with Philip Morris International and Selventa are at that front end of that candidate AOP. We've focused down on arterial stiffness, so we got a very clear kind of measurable endpoint at the end, and

we've got oxidative stress and some other things as the molecular-initiating event. And the thing I want to demonstrate with this really is that this is combining both clinical data and in vitro data to try and map out this whole pathway.

We've got a study just about to initiate with the University of Dresden in Germany, which will really hone down on this as an endpoint, to try and characterize this, both for cigarette smoke and for e-cigarette aerosols, and generate data that could then be put into this. Obviously this then should be chemical agnostic as the AOPs do. But this is work that's just started, but we hope to accelerate through this one quite quickly. We also have an AOP related to COPD as well.

Okay. So our approach is trying to combine the more traditional clinical efforts to characterize a new product with, I think, the more innovative approaches of omics and adverse outcomes pathway. And I think the two do complement.

There's one issue that I think, as a company who's trying to generate products that are substantially less risky than conventional cigarettes, fills us with difficulties, is the ability to characterize the risk of a new product and at the same time innovate those products. And we kind of see it in

the electronic cigarette arena, that the first electronic cigarette is nowhere like the -- that's currently here today, and that's over a period of 7 or 8 years. And there's a reason for that change.

So the first e-cigarettes weren't sufficient to be attractive to smokers to encourage them to quit smoking and move over to these products. So innovation, I believe, will remain a really key issue for next-generation tobacco and nicotine products.

So the dilemma then is, okay, if it's going to take us 2, 3 years and an awful lot of effort to quantify the potential risk of one prototype, and if we have done that for a cigalike e-cigarette, then actually we've already kind of missed the boat in terms of what we're going to do.

So I think there's a desperate need -- and I don't quite know how to do it, but I'm going to make a proposal here -- to find a way in which you bridge datasets through the things where you minorly modify the products, that it may be more attractive to smokers in terms of getting them to quit but haven't really changed the fundamental characteristics of the product. So you can tie yourself into databases.

And this is the proposal. So I think we've published what

we think is probably the approach to risk assessment of a modified risk tobacco product, which starts with product stability, looking at chemistry; you do your standard toxicology, you do biomarkers of exposure, and you try and quantify the uptake of the nicotine through PK studies. You then add to the adverse outcome pathway and system science. You do the biomarkers of potential harm in a longer-term clinical study, and then you do the population work on risk assessment. Eventually you do postmarket surveillance.

If that works, then there must be a way in which, when you do version two, as long as the version two isn't that different from a product component point of view and a chemistry and physical point of view, and it doesn't introduce any additional concerns from a population point of view, then there should in theory be no reason to have to do the 6- to 1-year big clinical studies again. At least that's what I'm advocating. But what you might need to do is use the adverse outcome pathway and the system science to help you provide your bridge.

So you've got some comfort that you aren't introducing any unknowns to this, even though you've got hopefully comfortable chemistry and the physical hasn't changed. And if that would be the case, then that allows innovation to continue much more

rapidly.

I think the consequences eventually of this, of course, are product standards. So if we look to the world of Swedish snus, for example -- and then I think we can tell if Swedish snus, from its characteristics as defined by the manufacturing and chemical standards, and our epidemiology is on Swedish snus. So we've got a bit of a connection between a product standard and the like.

And I think eventually where we'll have to end up, certainly for electronic cigarettes -- this is where Europe has gone quite quickly -- is you try and develop standards which relate to emissions in particular but also in terms of manufacturing standards for electronic cigarettes. So they're setting in a framework where you would expect those to perform at a certain risk level and you can then say, well, if you don't perform to that standard, then you're really not allowed onto the marketplace.

So that was my gallop to where we're trying to go. I think it's through experience, we felt, that we've really got to expand out our ability to collect data on next-generation products, and I'm seeing next-generation products evolve so quickly that I need to be flexible in terms of being able to

collect data in a quickly evolving world.

And I think that adverse outcomes pathways and using omics approaches can help us in that. And I think certainly for myself, who has to steward the products through in our world, and hopefully the regulators trying to assess what they do with these products, having some thoughts to how you do -- you bridge from one dataset to another dataset is going to be an important thing.

Just to acknowledge a broad range of groups that we're working with in this area, both other tobacco companies but also providers of research and also the people that really kind of move us forward, like Selventa, on the adverse outcomes pathway and the references related to this presentation.

Thank you.

(Applause.)

DR. DRESLER: Okay, our next speaker, I will say that I have not met -- yes, you are? Okay. So Dr. Emily Vogtmann from the National Cancer Institute will be speaking on Tobacco Use and the Human Microbiome.

DR. VOGTMANN: Before I start, I just wanted to thank the organizers for inviting me to come talk today.

In our group, we've been doing a lot of research on the

human microbiome. It's very much an emerging technology. We focus a lot on the methods for how to adequately collect samples, and then also looking at correlations between the microbiome and risk factors for cancer, and then looking at cancer as an outcome also. Clearly, for the talk today, I'll focus on looking at the associations between the microbiome and one of the risk factors for cancer. To start, I have no conflicts of interest.

So actually, Stephen gave a nice little introduction to NHANES a bit in his presentation, but I'm going to organize my presentation based on some study aims for a project we had funded through the FDA/NIH/Center for Tobacco Products money.

In this study, we aim to first look at the association between tobacco and oral health using the 2009 to 2012 NHANES cycles. Then we investigated the association between types of tobacco use and the oral microbiome, and then to look at the association between the oral microbiome and tobacco-related diseases, primarily periodontal disease at this point, and then to investigate the degree to which the oral microbiome may mediate the adverse health effects of tobacco use. So this is particularly where we're investigating whether it may be a biomarker of potential harm.

Now, this project is still in progress, so I'll be addressing these aims in different ways, but some of the data is not actually from that study. So to start, I'll talk about what I've done to characterize the association between tobacco and oral health in the 2009-2012 NHANES.

But before I get into that, you may be wondering why I'm interested in oral health, particularly while I'm at the National Cancer Institute. Some of the work in our group and others have found that tooth loss, which actually is commonly caused by periodontal disease, and also periodontal disease are associated with a number of cancers, including the cancers I have listed here. So that's where some of the interest in oral health kind of started for us at the NCI.

Also, as it relates to the microbiome, there are a number of pathogenic bacteria that are associated with periodontal disease. So that's kind of how we started linking the microbiome with cancer.

Okay. So in NHANES, it's a very large study that's conducted every 2 years, so there are cycles of 2 years that are representative of the U.S. population. One of the really neat things about NHANES is it has a number of, you know, questionnaire data, but it also has these mobile examination

centers, or MECs, and they do in-person assessments of the participants. So they weigh them and they collect blood. So then there's all sorts of biomarker data that's available. But they also have a direct periodontal assessment.

So a lot of studies have self-reported periodontal disease. This is an actual measured periodontal disease.

So in this study, we wanted to then look using the different types of tobacco that they assess on the questionnaire and then to look at how that may impact periodontal disease prevalence in the entire United States.

So to start off with, this is the first group we looked at. They are never tobacco users. So they reported, in the past 5 days, not using any tobacco product that was requested, which included cigarettes, cigars, pipes, chewing tobacco, and snuff. But they did have environmental tobacco exposure, which was assessed as either self-reporting having a smoker in the home or having detectable serum cotinine levels. And all of these were compared to never tobacco users and no environmental tobacco exposure.

So in this group, 15.5% of the 20.7 million people in this group that had periodontal, the 15.5 were associated with tobacco use.

When we looked at former cigarette smokers, in NHANES they ask about cigarette smoking in the much longer term, not just the 5 days. So these people reported being a previous cigarette smoker and having quit greater than 1 year since the questionnaire, and then with or without environmental tobacco exposure. You can see, without environmental tobacco exposure and with environmental tobacco exposure, less than a third was due to tobacco in those groups.

But now when we look at people who are currently using tobacco, so either the current non-cigarette tobacco users -- so these are the pipe, cigar, chewing tobacco, and snuff users -- a much larger proportion of this group had periodontal disease due to tobacco. And then the current cigarette smokers had an even larger proportion due to tobacco use. And overall, we calculated that 16 million cases of periodontal disease in the United States was due to tobacco.

So now thinking back to the microbiome and what I was talking about with maybe is the microbiome involved here, is the microbiome a biomarker of tobacco exposure, or is it a biomarker of potential harm? So we know that tobacco use is associated with periodontal disease, but then we also know that specific microbes -- I've listed one here -- are also

associated with periodontal disease.

So the hypothesis is that the microbiome may actually mediate the association between tobacco and periodontal disease and is then potentially a biomarker of potential harm.

So moving on, I'm next going to talk about when we looked at the association between tobacco and the oral microbiome.

Now, I've already been discussing the microbiome quite a bit, but I wanted to just kind of get us all on the same page about what it is. This is a tree of life proposed by Charles Darwin many years ago. Science has changed a little bit since then, and now with genetic sequencing, we can further characterize the tree of life more than Charles Darwin was capable of.

As people, we represent a very small portion of the tree of life, and microbes make up much of the rest of the tree of life. So microbes encompass bacteria, archaea, fungi, protists, and then viruses, which will not be found on the tree of life because they require another living being to survive. But they're also included in the microbiome.

And then, specifically when I talk about the microbiome, I'm talking about the microbial genes. So all of the work that we do is sequence based.

And then quickly, we as people are basically ecosystems. We have microbes living in and on our body in all different parts. And if you just look at these pie charts, based on the different colors that are found, you can see that the different parts of the body are very different. So the hair is very different from the colon, and you know, the esophagus is quite different from the skin.

Now, for all of my work that I'm talking about right now, I'm really only talking about the oral microbiome. But it will be important, going forward, to also look at how tobacco and cigarette smoking may affect other areas of the human body.

One of the really exciting things, too, about the microbiome is that it's able to be manipulated. When a person, for example, takes antibiotics, it dramatically changes their gut microbiome. You can also look at this pie chart. If you look up here, *Helicobacter pylori,* which is a known cause of gastric cancer, dramatically alters the microbiome in the stomach. So that's why we were interested in looking at how tobacco itself may affect the microbiome in the oral cavity.

Now, I did mention at the beginning our NHANES study is still under way, and I'll be really excited to share more of that data of how the different types of tobacco impact the

microbiome. But today I have data from a previous study that we conducted which looked at just cigarette smoking and then the upper gastrointestinal microbiome.

We conducted this study within a Cytology Sampling Study 2 population. This was a study conducted to look at different samplings for esophageal cancer screening. The study was conducted in Linxian, China, in 2002. For our microbiome component for this study, we only included males from the study population since in Linxian there were only, I think, two females total that smoked in our study. So we just excluded females in general since they are also likely gender confounding. And then in this population I present the cigarette smoking history because there were quite a few current smokers, almost 47%, and many fewer former smokers and then about 41% were never smokers.

And then this is how we collected those upper gastrointestinal tract biological samples. Essentially, this was for the screening study. They were testing these two different balloons. So this is Balloon 1 and Balloon 2, and what they were doing is the deflated balloon was inserted into the person's esophagus, and then once it was in there, they inflated it and pulled it back out to get all of the cells

lining the esophagus. But we call it the UGI microbiome because it's also going to have saliva on it. So it's really representing the upper gastrointestinal tract. And then we measured the microbiome using HOMIM array. It's a 16S RNA array itself. There's more details in the paper if you're interested in this.

Okay. So one of our big primary outcomes was just looking at a simple measure of diversity within the different smoking groups, so unique counts of different bacterial species or genera. And you can see that the current smokers appear to have slightly higher levels of both species and genera in their UGI samples.

When we put this into a polytomous logistic regression model, we found that this was statistically different from the never smokers. It did look like the former smokers were not different from the never smokers in this analysis.

We then restricted our analysis to current and former smokers to see if there were any associations between pack-years of smoking, and we didn't detect any in this population, but our sample size was small. So the NHANES study will provide an opportunity to look at that in greater detail.

And then when we looked at just former smokers, but that

was a really small group, there was no association in this study with years since quitting.

We then wanted to look also to see if there were any specific bacteria that were uniquely related to smoking status, and the current smokers tended to have a higher prevalence of the *Dialister invisus* bacteria and also the *Megasphaera micronuciformis* bacteria. These were both statistically significantly higher than the never smokers. Both of these bacteria are gram-negative anaerobic bacteria. So one hypothesis might be that these type of bacteria have a competitive advantage in a smoking environment.

So when the person is smoking, these bacteria are better able to colonize the UGI tract. Or one study did find that the *Dialister* genera and the *Megasphaera* -- or the *Dialister* genus and the *Megasphaera* genus were both present in cigarettes, not in Chinese cigarettes but in a study of cigarettes.

And so it's possible that introduction of bacteria could be possible through smoking, although it's unknown whether these bacteria are even viable that are in the cigarettes, and whether by smoking a cigarette with bacteria in it would actually be possible to colonize that tract. So this is certainly an important area to consider in the future.

When you look at the other studies that have been done looking at tobacco and microbiome -- there's a few, there's not a lot -- but there's no consensus on the association. Some of the studies have found that there are more unique bacteria in current smokers. Some of them have found that there are less unique bacteria in current smokers. And then some of them have found no differences. So it's really just -- it's unclear what's actually happening in these different smoking categories.

Similar to ours, most of the previous studies have had relatively small sample sizes.

About a week ago a study was published using samples from PLCO and the American Cancer Society samples, and they had a much larger sample size, but they didn't actually address, to my knowledge, this unique bacteria. They did find differences between the smoking groups, though.

And then cigarette smoking itself has been studied much more than the other types of tobacco. So there has been a lot of talk earlier today about, you know, the different types of tobacco and how they may affect some of these biomarkers. And most of the studies have only looked at cigarette smoking, to date. But again, we're planning on looking at the different

types that were addressed in NHANES in our future study.

So again, considering whether the microbiome is a biomarker of tobacco exposure or more likely a biomarker of potential harm, it does look like tobacco use does modify the microbiome. It's unclear how it's actually modifying it currently. We really need larger studies to see, you know, what's truly happening.

And it's also going to be important to understand the mechanistic relationship. I presented a couple hypotheses we had from our study, but it's really unclear what's actually happening and what the mode of action is in this case.

And so then, finally, I wanted to talk a little bit about kind of some of our future directions and what we're planning on doing to incorporate the microbiome and tobacco data with periodontal disease.

So in the NHANES study, we plan to essentially do what I just presented for the CSS2 population, but we want to do it in a much larger sample of NHANES. And also it's a U.S. population. So there's always possibilities that some of these things just differ by where the study was conducted because of different exposures.

We also want to determine the association specifically

between the oral microbiome and periodontal disease because the previous research with periodontal disease has generally focused on specific bacteria. And, you know, like we saw the *H. pylori* stomach, it's possible that once you have these periodontal pathogens, the rest of the bacteria changes also. So we're planning on looking to see if there's anything along those lines.

And then we really want to look to see if there is mediation of the oral microbiome between tobacco and periodontal disease, to see if it really is maybe a marker of potential harm.

So today I have really presented a little bit about how tobacco is related to periodontal disease. It has a large impact on periodontal disease in the United States. And then we also know from previous research that periodontal disease is associated with specific cancers.

Now, this data leads us to then look at the microbiome and, you know, how tobacco affects the microbiome, and then looking at how the oral microbiome and periodontal disease are related. And, you know, some really important work needs to be done to also see how the microbiome may independently be associated with cancer. A number of groups are planning these

types of studies. We're also starting to look at how the oral microbiome is associated with cancer. And, you know, I think this talk will be very different in the next year, 2 years, because this is, like I said at the beginning, a very rapidly developing field, and I think it's a really exciting area of research.

So I, like everyone, need to thank my collaborators, particularly my mentors, Christian Abnet and Neal Freedman. Both of them have been the ones that actually got the grant to do this work; and then all of the participants in CSS2 and our collaborators in China, and to thank the FDA Center for Tobacco Products. So thank you very much.

(Applause.)

DR. DRESLER: Thank you. We were scheduled for a break now, so let's just take a 10-minute break, and then we'll have two presentations and then the panel, okay? So that will give us an opportunity after lunch to stand up and stretch and maybe pop out to the kiosk. So 10 minutes, which will bring us right until 2:20, okay?

(Off the record at 2:10 p.m.)

(On the record at 2:20 p.m.)

DR. DRESLER: So our next speaker is Dr. Prasad from RAI

Services Company, and he will be speaking on Biomarkers of Potential Harm: Tools to Differentiate Tobacco Product Classes.

DR. PRASAD: So before I begin, I'll thank the CTP and FDA for giving me this opportunity to present some of our ongoing research.

By way of disclosures, I am a full-time employee of RAI Services, and the work is funded by R.J. Reynolds Tobacco.

Currently, there are several classes of tobacco products in the marketplace, and we can group them broadly into combustible and noncombustible tobacco product classes.

Use of any tobacco product comes with a certain degree of risk, and there is no tobacco product that's absolutely safe. So with the nonuse of tobacco being the safest option and cigarette smoking being the most harmful way of consuming tobacco, where do the noncombustible tobacco products fit in?

Within the broad spectrum of risk, a continuum of risk, epidemiology indicates that noncombustible tobacco products are less harmful compared to smoking. However, are there any short-term biomarkers or any other short-term tools available to differentiate these product classes? And also we need to remember that relative to smoking and smoking-related biology, there's limited information on the biology or biological

effects, the physiological effects of noncombustible tobacco products. And can we leverage any of the new and emerging technologies to differentiate the harm from tobacco products, these classes of tobacco products?

So that is the main objective of my presentation. I will share some data from some of the work, what we did from a clinical study, and some in vitro datasets and data from alternative model organisms, what we have been working with.

Let's begin with the biomarkers. And this is data derived from a clinical biomarker cross-sectional study, and I'm primarily focusing on biomarkers of effect. I just need to find the -- okay, biomarkers of effect. And these are primarily biological effects, what I'm looking at. And we covered this part.

In this biomarker study, the goal was to compare the effects of combustible tobacco products or cigarettes and noncombustible tobacco using moist snuff as representative product for the class. And this is not a brand-specific study. We looked at products, all different products in the product classes, combustible versus noncombustible. This allows us to look at the relationship between epidemiology and biomarker datasets.

In the cross-sectional study, we had 40 subjects, all male, generally healthy and -- excuse me. The cohorts, I abbreviated them. Smokers says SMK and moist snuff consumers says MSC. NTC is the non-tobacco consumers. I may be using these acronyms as we go forward.

In this study, we collected 24-hour urine samples under ambulatory conditions, and we also collected several samples for various experiments. We did a large number of -- we evaluated a large number of biomarkers of exposure and looked at a large suite of biomarkers of effect -- these are targeted -- and smoking-related biomarkers. Then we also performed several omic experiments, and I'll try to give you high-level results.

Just as a way of information, we published much of the data before and presented much of the data at various conferences, and we're in the process of publication, publishing the remainder. Some data are published already.

So in this cohort, we have our moist snuff consumers had higher exposure to nicotine and nicotine metabolites as well as higher exposure to TSNAs. Shown as an example, the NNK should have been NNAL, higher exposure compared to smokers. However, the exposure to noncombustible -- exposure to combustion-

related biomarkers was comparable between the MSC and NTC. Smokers had higher exposure.

We looked at a large number of smoking-related biomarkers. I'm presenting data from only those that differed in any of the cohort comparisons. We heard much about the blood cells. I'm not going to repeat through the blood cell counts and the cell leukocyte subtypes and a couple of lipoprotein-related markers, oxidized LDL and apolipoprotein B100. These were higher in smokers, and there was no difference between MSC and NTC cohorts. However, fibrinogen was lower in MSC.

Continuing on biomarkers of effect, we looked at a large number of -- a fairly large number of eicosanoids or arachidonic acid metabolites, which we heard in previous presentations. In this, generally, smokers had higher -- this black line in the spider graph. The smokers had higher levels of these biomarkers of potential effect -- biomarkers of effect. These include leukotriene E4, iPF2alpha, and isoprostanes and thromboxane metabolites, with the exception of iPFalpha-VI, which was higher in smokers and moist snuff consumers. Then we also looked at PBMCs, peripheral blood lymphocytes, mononuclear cells, a little bit more closely. We isolated PBMCs from the subjects, and we found that PBMCs were

higher, not a surprise in smokers. And similarly, T cells also were higher in smokers, but no difference between MSC and NTC cohorts.

However, the natural killer cells gave a different picture. Smokers had the least number of natural killer cells, followed by MSC compared to NTC cohorts. This suggests that all the inflammation, chronic inflammation, they were higher in smokers. NK cell function may be altered between smokers -- in smokers and moist snuff consumers. We'll come to this NK cell functioning and regulation a little bit later in the presentation.

We did several global profiling studies. I did metabolomics and looked at inflammation proteins and then transcriptomics from PBMCs and the epigenetics from buccal cells as well. We did it with the global metabolomic profile datasets. We did metabolomic profiling, using Metabolon technology, from plasma, saliva, and urine. And most statistical differences were between smokers, such as moist snuff, and moist snuff and NTC had least differences.

In the random forest analysis we performed, smokers could accurately be predicted in the prediction model, with or without the inclusion of nicotine metabolites. The prediction

accuracy, however, decreases for MSC and NTC cohorts if we excluded nicotine metabolites. It suggests that nicotine metabolites, maybe they are one of the key determinants for the separation in this model.

We also looked at two pathways more closely in the study, oxidative stress and inflammation. Again, smokers exhibited higher markers that indicated elevated oxidative stress, which mapped to vitamin metabolism, particularly vitamin C and vitamin E, and heme degradation pathways. Inflammation was also elevated in smokers, particularly purine degradation, which was prominent in smokers and less pronounced in moist snuff consumers. Purine degradation pathways were implicated previously in periodontal disease subjects. And we noticed several other differences in other metabolic pathways, which we'll publish soon.

We looked at inflammation proteins using a panel of inflammation proteins available from Myriad RBM. We used Luminex technology. Several proteins were consistently elevated in smokers, but fewer differences were found in moist snuff and NTC. The example, the soluble ICAM and MMP-9 were elevated in smokers, and we found only four differences between MSC and NTC. These are listed here.

Moving into transcriptomic datasets we did from blood PBMCs isolated from study subjects, most differences were found between smokers and in non-tobacco consumers for between MSC and SMK. However, no differences were detected between MSC and NTC under the analysis conditions.

We found that between these two comparisons, there were several genes that showed concordant expression by added data, which are listed in this next slide I don't expect you to read. But if you look at it, many of the genes have been reported for the smokers to be elevated and some were down-regulated also.

If you look a little bit more closely, the top hits are GPR15 and LRRN3, described in previous publications. We also have alterations in NK cell function and purines that are involved in lysing target cells. We'll come back to that later. And these preliminary bioinformatic analyses show that immune regulation is impacted in smokers. And we also have information evidence that prostaglandin metabolism is also impacted.

Then, because this was added data, we wanted to confirm it. So we took these 20 genes and then added a few more based on the biological significance, although they did not show concordant gene expression between the two cohorts we looked at

and -- okay, I went too fast. Let me go back. I looked at these gene expressions, and again, we could reproduce the gene expression added datasets in the RT-qPCR experiment.

We also found that AHRR, aryl-hydrocarbon receptor, we heard earlier, was elevated in smokers but not in moist snuff and non-tobacco consumers. The genes that were up-regulated and down-regulated were consistent between MSC and NTC. This led us to the conclusion, or this suggests that smokers exhibited distinct gene expression pattern relative to NTC and MSC. And the later two cohorts resemble largely each other.

Moving into methylation datasets, we took buccal cells and did methylation profiling using Illumina chips. Again, smoker cohort exhibits more profound qualitative and quantitative methylation differences. If you look at the top hits, AHRR was one of the top hits. Six of the seven loci were hypomethylated in this. And several of these genes, which you have seen from Dr. Bell's presentation, and some of the names probably are familiar -- I'm not going to go through the details. And methylation patterns in smokers were generally consistent with published data, published literature.

What we did was we took the differentially methylated gene loci and then grouped them based on the methylation levels into

three different signatures. One was combustible related tobacco signature. Another was tobacco related signature. The third one was the moist snuff related signature, just based on the p-values and the methylation levels.

Combustible tobacco related signature, or CTRS, is primarily derived from smokers. And moist snuff related signature is definitely from moist snuff related -- moist snuff consumers. And tobacco related signature had overlapped between the two.

We took the top 20 loci and then constructed a -- hierarchical clustering model, and smokers clearly segregate from the MSC and NTC cohorts. In a way, this is not surprising because 19 of the top 20 loci are derived from combustible-related signatures, so this is to be expected. This again suggests that top changes profiling, what we found, were derived from smokers. And much further bio-characterization by bioinformatic analyses are in progress.

So with this I would like to transit into some of the work, what we did with the in vitro and ex vivo models in this area. We focused again on inflammatory responses in PBMCs.

We know that smoking causes chronic inflammation, which yet it suppresses immune responses. That's been well

documented in the literature. There's a lot of data to show that smokers are at increased risk for inflammation, microbial infections, and various malignancies.

This is a hypothesis that can be tested in an in vitro model using PBMCs and toll-like receptor signaling. Basically what happens is, in toll-like receptor signaling, if toll-like receptors are stimulated with TLR-4 or TLR-3 using the ligands that are shown here, they secrete cytokines. If these cells are pretreated, the PBMCs are treated with the aqueous extracts of smoke, cigarette smoke, there's attenuation of these functions. We can look at cytolytic cytokine secretion as well as cytolytic functions of the PBMCs. Again, this again comes back -- cytolytic function comes back to CTLs and NK cells.

So I'm showing you one example. What we call this is whole smoke condition, meaning what we prepared from 3R4F cigarettes, and we characterized a little bit in terms of the constituents and normalized them for the equi-nicotine units of dosing and treated the PBMCs and looked for cytokine secretions. We looked at several of them, but they are shown with the TNF. And then as we increase the dose, the cytokine secretion is pretty much attenuated. But as the cells were treated under these conditions with nicotine, we see a

suppression only at higher doses.

A similar thing for the cytolytic functions. This is again mediated by NK cells and CTLs and PBMCs. Here we take K562 cells of the targets and look for cell killing. Pretreatment of PBMCs with the whole smoke condition medium inhibits the cytolytic function, whereas nicotine under these conditions did not.

So there's a lot more work that is going on in that area, and some of it was published. We'll get back to that. We're still working on that model.

Now, moving to the last piece of the datasets-wise, it's alternative models. I'm going to talk to you a little bit about the zebrafish and its potential use in the tobacco effect, investigating tobacco effects.

So there are a lot of alternative model organisms. There's definitely advantages of using these alternative model organisms. First of all, they're live animals we can follow through the life cycle, and they can bridge in vitro and in vivo models, and they can also help us in reducing animal testing.

We used previously two models. One was *C. elegans*. We published it, and I'm not going to talk to you about that

today, and I want to focus on zebrafish.

Zebrafish model presents us with a lot of strengths, several strengths. Primarily, they're easy to maintain. We can do high throughput, and several of the biological pathways are consoled between mammals and the fish, and behavioral tools available and we can do genetic manipulation. And shown here are the -- we can take the embryos, and we have the larvae by 72 hours.

So using this model, what we did was we looked at the effect of TPM on zebrafish. We took TPM and treated at 3 hours post-fertilization at four different doses and then assessed several functional parameters. And those that are shown in the red boxes are shown as examples in the next slide, what happens.

Here we have the TPM effect results in pericardial edema, as shown in the top panel, and that increases hemorrhage, as shown here, at the high doses. We used two doses, 0.4 nicotine, 0.4 µg of nicotine units and here again 1.4 µg. Effects are seen at the higher dose. And also we can see that it affects angiogenesis.

So a word about alternative models. There's much work that needs to be done on these alternative models prior to

getting -- putting them into toxicological evaluations.

Also, the cardiovascular effects reported here are not unique to TPM. Other toxins also exert similar effects. All the mechanisms could differ. Yet, the models are really potentially valuable.

In the last couple of minutes, I would like to get into the summary and conclusions. And we talked about multiple approaches to differentiate effects of tobacco products. First, we began with the smoking-related biomarkers, and these established smoking-related biomarkers are able to distinguish and discriminate smokers and moist snuff consumers. These are inflammation markers from blood cells to prostaglandin and eicosanoid markers.

And we also looked at several biomarkers of effect through different global profiling technologies, starting from metabolites all the way back up to the DNA. And we showed the differences, but like any other global profiling technologies, in any case we need to be further qualifying these so-called hits and the targets and the pathways in the follow-up studies. Otherwise, you know, we get into problems.

In general, our findings, what we found with the smokers, have been -- we are able to reproduce what was presented in the

literature. At all levels, our data was reproducible for smokers.

And the biomarkers of effect and the underlying biology, what we reported -- we didn't get into a lot of biological discussions today -- regarding MSC use, are really novel. Just to give you an example, we gave you -- presented more detailed metabolome from moist snuff consumers. We reported global profiling from inflammation from the buccal cells for the first time. I mean, there are a lot of novelties in this, and we believe that this information should build and guide us in regulation of differentiating for classes in the future.

In conclusion, these ex vivo/in vitro, and the clinical studies and the model organisms should be useful in evaluating different classes of tobacco products. Each system presents distinct advantages, and these are all mutually complementary. I don't see them as exclusive.

And overall, these biomarker discovery initiatives reveal that smokers exhibit more pronounced perturbations -- this is consistent with the epidemiology -- and the effects of what we see in smokers are distinct from moist snuff consumers and NTC, non-tobacco consumers.

And these data, what we presented from the biomarker

studies, allow us to say that moist snuff consumers more closely resemble NTC than smokers at this point, and we use this on generally healthy tobacco consumers. We want to look at the effects of tobacco. What does tobacco do to biology before it progresses to disease state?

With this, I just want to acknowledge our -- our references are there, but I want to acknowledge the contributions from my colleagues and collaborators. This is a team effort. The funding is provided by RJRT.

Thank you.

(Applause.)

DR. DRESLER: Okay, our next speaker is Dr. Shashi Amur from the FDA's CDER, Center for Drug Evaluation and Research, and she will be speaking on Biomarker Qualification at CDER.

DR. AMUR: Thank you. I have to lower this a little. Thank you. I would like to thank the organizers for inviting me for this workshop, and also I'd like to thank all of you for sticking around until the bitter end. I really appreciate some faces in the audience that I know. There are a lot of people who are attending through the WebEx.

I'd like to start with the statement that I don't have any conflict of interest to declare.

And here's a quick overview. I'd like to talk to you a little bit about our BEST resource, which explains biomarker terminology. Then I want to tell you a little bit about biomarkers in drug development, and then come to biomarker qualification, and end with some take-home points.

So BEST resource. It's a glossary of harmonized terminology for biomarkers and endpoints. BEST stands for Biomarkers, Endpoints, and other Tools. I'm part of the FDA-NIH Biomarker Working Group who has been working a lot on these harmonized terminologies, and we are continuing to work with that, and I'll explain a little bit about where we are going from here.

The first question people ask is we know what biomarkers -- we know about these terminologies. Why do we need any kind of harmonized terminology? One of the problems is when somebody says prognostic biomarker, for example, it may mean something to one person and something else to another. For example, prognostic biomarker, to somebody, might mean something which will tell you how the disease is going to progress, and for others it may be that and also susceptibility biomarkers combined.

So, you know, it's good to have terms or terminologies

that are clear and consistent, and that will really help the science as well as science progress and help us in the regulatory side.

Here is one slide to tell you why a glossary is needed. There are so many terminologies, and some people use those interchangeably. For example, fit for purpose. And then we have context of use. We have qualification. So there are a lot of terminologies which people use. Analytical validation, for example, versus clinical validation. Then there's qualification. So it is an important thing.

I'm not going to go into the details, but I just want to tell you that currently we have identified that these are the biomarker categories. This will be in a draft form. It will be available to you to see it on a website, and then we're going to wait for public comments. These are by no means final, but these are the ones we all came up with as a group, and it's been going on for several months.

So one is the susceptibility/risk biomarker. Second is the diagnostic biomarker; prognostic biomarker; predictive; monitoring biomarker; pharmacodynamic or response biomarkers, which indicate what we call the efficacy biomarkers also, such as surrogate endpoint biomarkers; and then there are these

safety biomarkers.

People have been talking a lot about biomarkers in drug development, and we view, in FDA, biomarkers as -- especially in CDER, we view that as a drug development tool. The short form for drug development tool is DDT, not to be confused with the pesticide.

Biomarkers have been used for many, many years in drug development. You can use biomarkers in basic research and to understand disease pathways. You can use it in drug discovery to understand mechanism of action of therapeutics and also to support drug target selection. You can use in preclinical development through safety assessment, mechanism of action, dose determination. These are only some of the uses of the biomarker. There are many more.

In clinical development, it can be used in patient stratification, selection, enrichment, dose selection, safety assessment, efficacy assessment, and also in some cases as surrogate endpoints.

There are two main pathways to integrate biomarkers in drug development at CDER FDA. We've been often accused that for us in CDER, CDER is the world, and we use FDA a little loosely. So pardon me for those transgressions.

So there are two pathways. One pathway to integrate biomarkers in drug development is through acceptance of the biomarker through IND, NDA, and BLA submissions, which is the drug approval process. The second pathway that we have is biomarker qualification.

One question that was asked -- that's asked often is now that you have a biomarker qualification pathway, can we still go ahead and use biomarkers through the acceptance to the individual drug submissions or biologic submissions? The answer is yes, these are two separate pathways, and they have their advantages.

If your objective is to use the biomarker in a single drug development program, obviously you should go through the acceptance side of it, through IND, NDA, and BLA submissions, where the sponsor contacts the review division, talks to the review division, and convinces them this is how we want to use the biomarker and why we want to use the biomarker, and the biomarker gets accepted into drug development.

If the objective, on the other hand, is to establish the biomarker for use in multiple drug development programs, biomarker qualification is your pathway of choice.

Any individual or group can send in the submission.

However, in our experience, generally consortia are the ones who send in submissions to the biomarker qualification program. I think the reason is because the risk as well as resources are shared among consortia members. The advantage here is once qualified, the qualified biomarkers are announced as draft guidance and eventually as final guidance, so the information becomes publicly available and easily available.

So I've been talking about biomarker qualification but have not defined it yet. So biomarker qualification is a conclusion that within a carefully and specifically stated context of use, the biomarker has been demonstrated to reliably support a specified manner of interpretation and application in drug development.

So an often asked question is, is this validation? Is this qualification? So we call it qualification because we are saying only for that particular context of use, we are saying this biomarker is good in drug development. So if it's not in the context of drug development, we do not consider these for the biomarker qualification program.

The term that's very important and is often used is context of use. Context of use is nothing but a statement that describes the manner and purpose of use, the how and why,

maybe, for the biomarker in drug development. And this context of use needs to be defined early because you need to have data to support the context of use in order to qualify a biomarker. So the context of use thereby, it is responsible for the level of evidence needed, so it drives the level of evidence, and the level of evidence itself drives the qualification process.

It would be wonderful to have evidentiary standards for different types of biomarkers for those, for there are multiple contexts of use. We are not there yet, but we're working towards that.

I would like to focus your attention on this bullet. We held an FDA co-sponsored workshop with M-CERSI -- that's the University of Maryland CERSI program -- and Critical Path Institute on "Evidentiary Considerations for Integration of Biomarkers in Drug Development." We discussed enrichment biomarkers and safety biomarkers.

Then I'd like to focus your attention on the last one, which is the workshop which is planned for April 14th and 15th of this year. It's an FDA-FNIH biomarker consortium workshop, and we will be discussing evidentiary criteria for safety biomarkers.

In the meantime, here are some of the considerations which

are general considerations for biomarker qualification. First, you need to understand the type of biomarker you have and the context of use that you have for that biomarker, proposed use for using a biomarker in drug development.

So the next thing that's very important is to characterize the relationships among the biomarker, the clinical outcome, not clinical endpoint, and the treatment, when applicable, that's required for the proposed context of use.

Another important consideration is the assay. We do need you to use an analytically validated method. We are not actually qualifying the assay; we are qualifying the biomarker. However, we need to be sure that the biomarker assay that's used is reproducible and accurate and we can believe the results that are coming off of the assay or the test.

It's important to understand the biological rationale, if known. I do understand -- I mean, I've been listening to you, and there's genomic signatures, there's proteomic signatures, miRNAs, where you may not understand the biological rationale, starting out anyways. And many never do. So that is totally understandable in those scenarios.

Another thing to consider is the type of data available. From our perspective, the retrospective data as well as

prospective data are welcome, and for certain types of biomarkers you could also use registry data. And randomized controlled trial data, of course, are pristine, so those are compared to the others. So those are definitely welcome.

Another important consideration is reproducibility of data, particularly -- I mean, I am a molecular biologist by training, so I know when I used to do the microarray experiments some time ago and you get the signature one time, and the next time you do an experiment the signature is very different or there are only a few that overlap. But through the MicroArray Quality Control Consortium and other efforts, there has been a lot of standardization. So reproducibility is improving, but that's something to keep in mind.

Sometimes we hear, we have published the results in one -- you know, this very good journal. Why do we need to repeat the data and repeat the experiment? And reproducibility is key, as we have found the hard way.

And as I mentioned, retrospective data are okay to use. However, you do need to consider using pre-specified statistical methods to avoid or minimize bias, to demonstrate the relationship for the context of use. The overall strength of evidence is what helps us decide on whether the biomarker is

qualifiable or not.

Oh, I'm going slow. I've been asked to address single versus composite biomarkers. It is possible to qualify either a single biomarker or a composite biomarker, which is referred to as panel of biomarkers to be qualified.

There are some considerations, though, if you are thinking of using composite biomarkers, pre-analytical considerations, stability of individual biomarkers. Does the sample collection time matter? And analytical validation again. Sorry for repeating, but it's important. Factors that influence the variability of individual biomarker levels. Whether to normalize the biomarker levels or not.

Weighting. Should it be an equal weighting or an unequal weighting? And if either answer, why? And assessing variability of the composite measure and understanding of factors that influence this composite measure are important. Again, you need to understand how it's going to play out into decision making in drug development. So you may need to identify a score or a threshold score where, you know, above which or below which it may be significant.

I'm not going to go into the details, but here are the three stages in the biomarker qualification process:

initiation, consultation and advice, and review.

So far, we have qualified 13 biomarkers. Ten are nonclinical ones, in white. The ones in light blue-gray there are the clinical biomarkers, mostly for patient selection, so far. But we do have 25 submissions, active projects ongoing, as this number indicates; 4 are in the initiation stage, 18 are in the consultation and advice stage, 3 are in the review stage.

If you look at what kind of biomarkers do we have, we have -- one-third of them are for patient selection. About half of them are for preclinical and clinical safety. As you could see, safety biomarkers are very important. Seventeen percent of them are for pharmacodynamic biomarkers or response biomarkers. There's one on patient compliance.

There are some obstacles to biomarker qualification, as you know. Lack of a clear context of use is one. Insufficient supportive data is another. Resources are always a big problem. Challenges in aggregation of requisite data. You may need permission to use somebody's data; that can be a problem. How the sample is stored. Whether your biomarker is stable or not. Assay validity/reproducibility. Lack of clear evidentiary standards that we are trying to address.

And here are some take-home points. So biomarkers can be integrated into drug development through either of the two pathways, one through the regulatory submissions for drug approval in the context of a single drug, or biomarker qualification.

Biomarker qualification is a voluntary process which is intended for biomarkers to be used in multiple drug development programs. Once qualified, the biomarker can be used for the specific context of use in regulatory submissions without having to reconsider and reaffirm or confirm its suitability. That's an advantage of biomarker qualification.

We are not charging any fees for evaluating biomarker qualification submissions.

And finally, I would like to encourage early engagement with FDA on biomarker qualification.

I'd like to thank Drs. Woodcock, Buckman-Garner, McCune, Leptak, Marianne Noone, Dr. Sanyal, and Dr. Haskins.

Thank you.

(Applause.)

DR. DRESLER: Could we please have the panelists come up to the front from this last session?

Okay, so questions from the microphone or send them in

online, or if you want to raise your hand, you could still get a card. And the panelists can actually ask each other questions, too.

So this was one of the questions that I had that makes me think back to the last session. When we're looking at these and looking at the methylation -- and Dr. Prasad, this was, I think, mostly with yours -- what cigarettes and what moist tobacco products were being used? And do we need to -- when we start looking at these very complex, detailed sort of metabolomics or proteomics or the gene arrays, whichever it is that has this plethora of data, do we need to know what the actual product was when we were talking about these outcomes?

DR. PRASAD: In this study we did as a discovery study --

DR. DRESLER: Um-hum.

DR. PRASAD: -- we focused on the product classes. Probably, as I said, whatever targets, what we found are hits that need to be qualified in an independent study either internally or externally.

Now, coming back to the specific branded products, if somebody wants to use a regulated for a submission, we probably may not have to do the entire profiling. I mean, maybe Dr. Spira may have a different view, but it will be simpler if

we can do it with the non-targets which have been established in the discovery phase. It's the same principle as what Dr. Hatsukami described.

At this stage, we start with the exploratory discovery research, and then as we progress through, we'll narrow down to fewer biomarkers, and then we can apply those fewer biomarkers in the submission process. That's my view.

Rather than doing eight for each product, if we have to do profiling of everything, we can do the profiling. It's not the difficult part, but putting together the information and the bioinformatic challenges and what it means for each product, is it necessary?

DR. SPIRA: I would agree. I mean, I think the example I showed today, we started off on a 22,000-microarray experiment -- 22,000 gene experiment and ultimately go down to 23 genes, and the final product that's now on the market is a 23-gene -- it's an assay measuring 23 genes. So inevitably, you do want to narrow it down.

Now, what platform you use to go after those, it's a whole other question. What's interesting, in the case of Veracyte, who's commercialized that test, they've stayed on a microarray platform to measure the 23 genes. So they're measuring a lot

of other things, but they're just focusing on 23 genes for their classifier.

DR. DRESLER: Okay. Yes.

DR. PROCTOR: But I think maybe the point to your question was do you need to be able to characterize the products is probably correct. And I know in the work we're trying to do on the electronic cigarettes, it's such a range of product that you would expect there to be a different consequence to some of those products. The same may be true for smokers. And I think also the one thing I learned from the last biomarkers of exposure, you also need to characterize the subjects. And then I came and told you -- well, I get excited when I come to these workshops. I came excited about the exposome the last time and trying to figure out, so what was your starting point when you got volunteers to use these things, because you're going to need to understand to figure out. So what were the responses that you got?

DR. DRESLER: I think it was you who actually used the gene-environment interaction.

DR. SPIRA: Correct, right.

DR. DRESLER: So that goes to you are going to have know the genetic plus the environment, which would be the product.

DR. SPIRA: Right. No, I think in general, certainly for a disease like lung cancer, it's clearly both genetic susceptibility combined with the exposure.

I had a question for Dr. Prasad related to sort of the methylation and the buccal mucosa. Have you looked at any gene expression changes in those cells? I know we've been challenged with getting good RNA out of the mouth, but have you look at the relationship between methylation and gene expression? And if not, I know there's been publications from other groups. I know of at least one where they biopsied the cheek in smokers and identified gene expression changes. Have you been able to look at the correlation between your methylation events and their gene expression changes?

DR. PRASAD: We have not been able to get good data for RNA from initial experiments from -- these are mouthwashes. We did not get good quality for RNA. This was done a few years ago. But the methylation data, first of all, it's been produced for many top hits of genes, what we looked at from other studies. I know that at least two other people have looked at this concordance between buccal cell DNA methylation and buccal cell gene expression. There has been some concordance.

And I think they also reported, if I'm not mistaken, AHRR is one of the things that they show in the gene expression. And there's a publication by Martin Schauer from Europe, from UK. They did the comparison between buccal and blood and then gene expression.

There was one more paper on snus from Sweden. I think they did a comparison between smokers' gene expression and methylation data. I don't know if -- I don't recall if all the genes, all the loci, all the differentially methylated genes and differential gene expression was concordant in those studies, but at least for some genes they have been able to demonstrate that. At least coming back to my data, what I presented, I didn't have time to go through that.

AHRR was hypomethylated in buccal cells. In blood PBMC gene expression, we showed elevated expression of AHRR that's concordant. In another publication, people have showed GPR15. I think Dr. Bell or somebody showed the GPR15 methylation, and GPR methylation we found as one of the hits in the buccal cells, but in the blood, GPR15 expression was elevated. It was not tissue to tissue, but in the -- we could demonstrate that.

DR. SPIRA: Thank you.

DR. DRESLER: This is a question that goes back to

Dr. Amur, something that you said. But this was a question that I had for several of the panels. You had talked about assay validation, and one of the things that we've heard in the past 2 days, there's a whole variety of assays in order to look at these biomarkers. Where are we with having this variety that we've heard in the past 2 days, with validation of the assays such that they give reproducible results? Because everybody's developing different assays. Anybody want to comment about where we are with just reproducibility of the different assays?

DR. AMUR: If you want to take it, I'd be happy -- okay.

DR. PRASAD: No, no, no. It's not mine.

(Laughter.)

DR. AMUR: No, it's a good question, and there are different types of biomarkers. I mean, blood pressure measurement. Blood pressure itself is a biomarker. Then you have imaging biomarkers, and then, of course, you have the blood-based biomarker, CSF, urine biomarkers. Those are the ones which people think of when you hear the term biomarker. So what you would have for imaging biomarkers will be very different from something for plasma fibrinogen. So I mean, we do understand that.

So I don't -- I mean, the CDRH would be the right office, or should I say Center, who can address your question directly. But from the biomarker qualification perspective, what we do is that we have a biomarker qualification review team. So we have a clinical reviewer, we have a statistician. We also usually have somebody from CDRH who helps us to evaluate if the assay is analytically validated.

So as I mentioned, for blood-based ones, it'll be different from the ones for imaging. How is the imaging captured? You know, what kind of software is used and how is it analyzed? How are you going to see something that's done in one center? Is it similar to what's happening across -- so all of those questions are asked up front and then, you know, we work with the submitters in a very collaborative way. It's a very scientific collaborative branch of FDA.

So we handle it on a case-by-case basis, but we are coming up with some kind of a white paper sometime soon, and then, you know, we'll be able to describe some of those with lessons learned. And the workshop that I mentioned to you, that's coming up in April, and they're also working -- there is a group of them who's working on analytical validation and considerations. So we'll know more as we go on. But we do

have experts on our review team.

DR. DRESLER: You know, it made me think of the doctor from CORESTA who spoke this morning. Dr. Prasad, the work that you were doing with it, it directly goes to this, how these tests are being done around the world. But I'm looking at the plethora of tests we've heard in the past 2 days and the timeline that I know the industry is interested in working on having those assays validated and just the reproducibility of the assay. The heck with whether it actually says what it's going to say. And I'm just wondering if you all want to comment on this. So yes, it needs to be done for it to be qualified, but do you have any suggestion of a timeline or thoughts about having assay validation and the standardization of those tests?

DR. PRASAD: I really don't know how to answer that. It's a huge effort. For many of the molecular markers, the way I'll at least pass on that I approach assay validation. First of all, we need to make sure that the result is reproducible with whatever biomarker or potential candidate we are measuring.

So we approach what at least we have implemented, is we'll try to do external validation using an independent set of samples, preferably, and in more than one different type of

populations and preferably done at more than one location. At least two or three people, if not a formal ring trial at a couple of different labs. That's one thing.

Many times what happens in the discovery studies at least, the sample becomes limiting and we may go for internal qualification. For example, that's what we did on our transcriptomic data. And if we have a sample, we definitely like to take it to external qualification. Those, to a large extent, should be able to say that the result is reproducible or not. For the biomarker, are we seeing a true result, or are we looking at a red herring, if you will? That, in fact, was one of my questions to Dr. Shiels earlier for the cytokine or inflammation proteins. What other alternative assays can be used?

DR. DRESLER: Um-hum.

DR. PRASAD: I know we all use Luminex technology. If you go back and do the same thing, do we do external qualification, or do we do the assay again because there are antibody-based assays? When it comes to the more established targeted assays, of course things like mass spec methods that I mentioned earlier. In a different context, a GLP type of assay should be developed.

The question is whether the consortium would develop or whether a new general company would develop or a lab would develop. These are the much larger questions. Ideally, we would like to have a consortium just as the PSTC consortium is in place and they develop -- they determine someone to work for qualifying some of the biomarkers, kidney biomarkers. It would be nice if we can get together -- all interested stakeholders get together and identify a few biomarkers. These are a couple of biomarkers, whether it's iPFalpha-III or whatever it is. If we qualify that, that would be a good effort.

And the challenge is some of the samples may be stored, they're stored in freezers. Some of the biological samples may not be as stable when time comes to take it out and try them. So these are all the challenges Dr. Hatsukami mentioned, and the IOM booklet also talks about these challenges. These are not particularly new.

DR. DRESLER: Dr. Proctor.

DR. PROCTOR: I think it's going to be a weight-of-evidence approach in this, and there are some very established things. The work of Drs. Hecht and Hatsukami in some of those traditional markers, it's very solid, and I think we can rely upon that. When we get to the newer work, it's far less

defined.

And I think what actually the consequence will be is that the products that you're evaluating will have to be very, very different from cigarettes. So they'll have to have a biological profile that the signal is pretty certain that you've got something going on, which is very different from a cigarette. Your challenge is going to be any kind of unknowns, I think. So because cigarettes cause so many biological consequences and they're usually pretty intense, I think something which is a much cleaner system is likely to be distinguishable, even when you haven't fully defined some of these techniques. And by weight of evidence, then hopefully you can get fairly close to being not completely certain but reasonably certain this is going to be a different -- in a different class, but really focusing down, okay?

But if this introduced anything unknown that we didn't have with cigarettes, that's, I think, more of a challenge, and that's where I think some of these broad approaches to see whether there are some signals you wouldn't expect are being seen are quite important.

DR. DRESLER: Dr. DeBord.

DR. DeBORD: Yeah, just even the cigarette smoking aside,

you know, this is one of the challenges that has been, you know, about biomarkers for 30 years that we've been using them, because human biomarkers first came upon the scene back in the 1980s or even sooner or even earlier, and everyone was really enthralled with the use of biomarkers and how it was going to change the world, and then they got disappointed really quickly. And that's basically because, you know, they weren't asking the right questions or were trying to answer questions that weren't necessarily going to be answered by that biomarker.

And so that's why we really pushed the validation, and why a lot of biomarkers aren't being used for risk assessment or for decision making is because they haven't been validated. And so there are different ways to validate biomarkers. You certainly have to validate them in the lab and make sure that they're reproducible in the lab. But then you also have to validate them for the population that you're going to be testing, at least in the occupational health world.

So we're a long way from that. Everyone wants to identify that new biomarker, but no one really wants to take the work to make it -- you know, to take it to the end to validate it so that it can be used in policy decisions.

DR. DRESLER: Sir.

MR. SENGUPTA: My question is for Dr. Amur. What is done right now once you get the applications for biomarkers on FDA's side?

DR. AMUR: Are you asking me about the process itself?

MR. SENGUPTA: Yeah.

DR. AMUR: Yeah, in the initiation stage -- as I explained, there are three stages, initiation, consultation and advice, and then finally the review stage. So in the initiation stage, all we get is a letter of intent. So what's the biomarker? What do you want to use it for? Some kind of a general plan as to what's known and what needs to be done. So maybe five, four or five pages or sometimes two pages.

So once that comes in, we have a go/no go decision. Basically, the program first looks at it and says, you know, this is worthy. Let them go to the next step. So the next step is we contact the office directors and say, you know, would you like to get a reviewer for our BQRT, or the biomarker qualification review team? Office directors can then look at it and say is this, you know, something valuable to them? Do they have the resources needed?

Then once that is cleared, we get a reviewer. And then

the BQRT also evaluates the letter of intent, and that's another go/no go kind of a thing. So once that clears it, it's accepted into the program. Then we send recommendations and suggestions as to what we want to see in the next step. That's when the consultation and advice starts. So once we get the document from you, we have a face-to-face meeting to discuss it. And when we think that the biomarker development is complete, we invite you to send in the final qualification package, and then we review that, and if we need any clarification, we again reach out to you. Like in NDA/BLA situation, they have the information requests. So we have -- we might make information requests.

So once we decide to qualify the biomarker, we issue a draft guidance. So that's publicly made available. The reviews that are associated with the qualification are also made available. There's an executive summary so you have all of those. After 60 days of public comment period, we then revise it, if needed, or update it, and then we issue a final guidance. So that's the process.

MR. SENGUPTA: Thank you. And I think I saw somewhere like fibrinogen is -- have you done something with fibrinogen?

DR. AMUR: Yes. We have qualified three clinical

biomarkers. One is galactomannan from serum and bronchoalveolar lavage fluid. And also then the plasma fibrinogen for COPD, in COPD, for enrichment, the patient enrichment in clinical trials. And then, finally, the total kidney volume, that's an emerging biomarker. That's also for enrichment in clinical trials. But this is for polycystic kidney disease.

MR. SENGUPTA: Thank you.

DR. AMUR: You can find all the information on our website, and please feel free to contact me.

DR. DRESLER: Okay, a high-level question for each of you, so prepare to answer this one, please. Given the large amount of data that's generated by all of these technologies, what are some of the considerations or approaches that need to be taken before applying them to any public health setting? So right, the Center for Tobacco Products makes their decisions based on a population public health basis. So a large amount of data. What are some of the considerations or approaches that need to be taken to look at that? Don't laugh at me.

DR. AMUR: I'll start from the other end. One of the things we have seen is data standards. People there have multiple datasets, and to combine them is challenging because

the data elements or data standards have not been defined. So if you have one dataset you want to combine with another, you don't know for sure if this header means exactly what the other header means. And so that's a big challenge. And in one of the examples of our qualification step, they took about 2 years to define all of the data elements and standards and to aggregate the data.

With the CDISC standards, I believe life has become a little easier from the CDER perspective, but that's still a challenge because a lot of different people call things in different ways, and there's usually no defined kind of a document available with every Excel sheet you see. So I think one of the challenges to me is data elements and data standards.

DR. DeBORD: For GWAS studies, genome-wide association studies, a lot of the journals have come together and set standards on what data should be collected, how it should be collected, and what the measurements should be and how those measurements should be reported.

I think NIEHS had a workshop a year ago in January, January 2015, in which they were trying to develop standards for EWAS studies, environmental-wide association studies, so

that way, you know, we could do better comparisons between studies and then be able to mine that data a little bit better because the measurements -- you know, the metrics were all the same. And so I think it's going to take something like that, you know, where we can cross over studies. But we also need better mining informatics.

You know, we have these large databases, and right now I don't know that there are good ways -- you know, there are still a lot of different approaches being applied to large databases on how do you mine, you know, like omics databases to get relevant information, because it takes a long time. I'm not really up on the informatics part of it, but that is a challenge is to -- you know, what is the relevant information?

At NIOSH I also work on sensor technologies. And so we actually, you know, have the same problem because we used to have sensors that would take a measurement every 15 minutes, and now we have a sensor that takes a measurement every 5 seconds. And so, you know, what is the relevant data? So big data, in itself, is a problem across the scientific industry.

DR. PROCTOR: I'm going to draw on that. I love the picture of Darwin's tree of life, and I think the recommendations don't always have to -- sometimes a simple

hypothesis is not that simple, but it's actually quite instructive, and I know we know it's a lot more complicated than that. Actually, that served a purpose.

And I think in this case it's almost back to one of the earlier presentations. Think about what we're trying to achieve here. So what we're trying to achieve, we've got a dominant tobacco product which causes a huge amount of disease, and we're trying to quantify what might be new products that cause much less disease.

So do some basic stuff. So if there are no carcinogens in the product, then okay. Well, we need to look at irritation that might happen, and there may be still some consequences. But let's set up the scientific hypothesis for what this product should be.

Now, it gets tricky in the context -- this is where Europe, I think, separates from where you all are here because in Europe we've taken that approach to electronic cigarettes by saying okay, here's a category, let's set some basic standards, and then evaluate those against a particular set of standards, they shouldn't have any more metals in this, they shouldn't have any more carbonyls in this, and let's do the evaluation, whereas here we're seeing that you're going to have do it

piecemeal, specific product by specific product. And I think that becomes a lot harder to do, particularly when we're in this kind of gray area where the research isn't absolutely obvious what you're going to do to characterize this. I do think the further away these new products are from the dominant product, the easier it's going to be.

But it's still going to be much trickier, unless we take a more holistic approach to, okay, what is a snus category? What is an e-cigarette category? How do we define those? And then we worry about population health. I think that's quite a separate thing from trying to find the actual risk characteristics of some of these categories at the beginning.

DR. EDWARDS: Okay. So I'll first echo the data standards and meta-data standards. One of the things that we're working on in the knowledge base effort is trying to tie things into biological ontologies on the metadata side. And so one of the things we're hoping to do in the not-too-distant future is have the components that are described in the knowledge base tagged with ontology terms. And then specifically, in my group, we're doing the same thing when we're doing these data mining exercises, is trying to map into ontology space to connect up the more bioinformatics-driven AOPs with the expert-driven

AOPs.

Now, as far as the data itself, a lot of what we've ended up doing, because we do end up merging a lot of disparate data sources, is we discretize. So we take whatever the standard is for a given dataset, turn it into yes/noes, and then move -- that's what moves into the metadata, I mean into the data mining analysis. And so I was loath to do that at first because I'm thinking we're losing all of this information, but if you don't have the needed information to interpret the continuous data, you may as well discretize it.

DR. VOGTMANN: One of the things for kind of going along with that and, you know, just combining datasets, at least for the microbiome, one of the big things that we've been pushing for is quality control samples. In the fall, NIST is hosting a workshop on generating some of these samples. But because some of the previous work, we've seen in a microbiome quality control project we conducted some lab-to-lab variability, mostly likely driven by DNA extraction. But if there were quality control samples that were consistent between experiments, there may be a way to extrapolate the data so that you don't have to dichotomize it or make it much more simple.

There may be a way to manipulate the data, you know, in

the future to know this drift or whatever is happening to be able to combine the datasets. Because I do think with all of these omic technologies, you know, and having multiple technologies for the same samples, you need larger and larger samples sizes because of all the multiple testing. So it will be really important in the future to have this ability to pool multiple datasets.

DR. DRESLER: Dr. Prasad, did you want to have the last comment on that? You don't have to, but if you would like.

DR. PRASAD: I'll try. Going back to what Dr. Proctor said, there are so many different products, and exposure is the one that we need to set right. So as the products evolve, looking at those new products may require -- should not become too cumbersome. My personal view is because we have so much information on cigarette smoking and the pathobiological effects and the large amounts of datasets available, whether gene expression or methylation or across different populations, so if we can use that as the baseline information -- and then I don't know how we probably can do that. If we can blend those datasets and then arrive at some few key pathways or some basic information, what we think we can use, for example, isoprostanes, AHRR, or some other pathways, and then blend with

one of the function assays and use that as a benchmark, this is the most extreme and no harm, what we have in this continuum.

How would you present your product? What information would be required? So that way we can stop doing -- generating whole amounts of gene expression datasets or methylation datasets. If we can focus on a few things and do them right, that may be one way perhaps that CTP consider for regulating the products in the future.

Because we have a lot of information on cigarettes and cigarette smoking, and there's still so much and it's amazing and everything else, we can use that as a baseline and use that to advance further regulation for the products or building science on the products as appropriate.

DR. DRESLER: All right. Any other comments? Thank you all for hanging in there. And no more questions. It's been a long 2 days, but you guys have been terrific. Thank you so very much for bringing up the last session and the panel discussion. So thank you so very much. And to all of the speakers also, thank you.

And then I also want to thank Cindy Chang and her team and the organizational team, Dhanya and her team, for organizing this. They always do such a terrific job. So thank you.

Thank you, everyone.

(Applause.)

(Whereupon, at 3:30 p.m., the workshop 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:

BIOMARKERS OF POTENTIAL HARM: A PUBLIC WORKSHOP

April 5, 2016

Silver Spring, 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.

_____

Timothy J. Atkinson

Official Reporter

Professional Video Associates, Inc.
2515 Saint George Way
Brookeville, MD 20833
301-924-1556