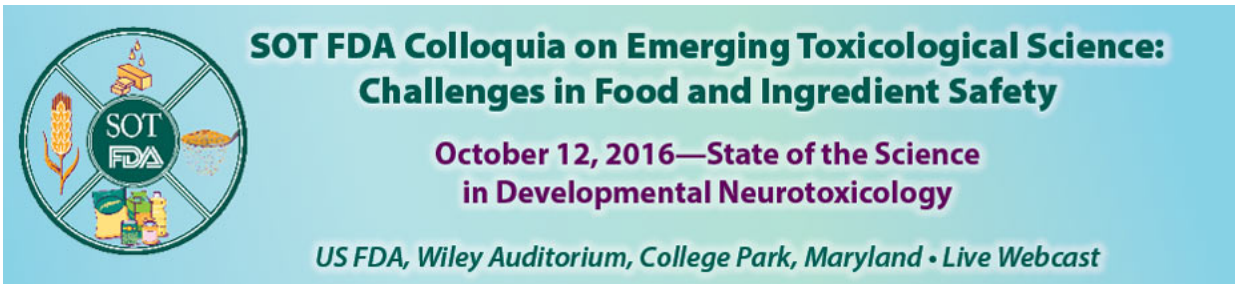

Real-Time Captioning

Note: This is not a transcript.

State of the Science in Developmental Neurotoxicology

Timothy J. Shafer, Chair, US EPA, Research Triangle Park, NC
Jeffrey J. Yourick, Co-chair, US FDA, Laurel, MD

8:00 AM–8:30 AM	Badge Pick Up
8:30 AM–8:40 AM	US FDA Welcome and Overview, Mickey Parish, PhD, Acting Director of the Senior Science Advisor Staff, CFSAN, US FDA, College Park, MD
8:40 AM–8:50 AM	Welcome from SOT and Introductions, Peter Goering, PhD, SOT President, US FDA, Silver Spring, MD Speaker Introductions, Timothy Shafer, Chair, US EPA, Research Triangle Park, NC
8:50 AM–9:30 AM	Developmental Neurotoxicity Testing: An Introduction to the State of the Science and Opportunities for Improvement Charles V. Vorhees, Cincinnati Children's Hospital Research Foundation, Cincinnati, OH
9:30 AM–10:10 AM	Zebrafish as an Alternative Species for Developmental Neurotoxicity Testing that can Provide Hazard Identification and Mechanistic Information Randall T. Peterson, Harvard University, Boston, MA
10:10 AM–10:30 AM	Break

10:30 AM–11:10 AM	*In Vitro* Approaches to Screening Compounds for Developmental Neurotoxicity Hazard Ellen Fritsche, University of Düsseldorf, Düsseldorf, Germany
11:10 AM–11:50 AM	Adverse Outcome Pathways for Developmental Neurotoxicity Anna K. Price, ECVAM, Ispra, Italy
11:50 AM–12:50 PM	Roundtable Discussion Timothy Shafer, EPA, Moderator Jason L. Aungst, FDA, Panelist All speakers
1:00 PM-2:00 PM	Informal Lunch for Speakers and FDA Employees CPK 1 Room 1B042 (Bring your own lunch) FDA employees are welcome up to the capacity of the room.

US FDA Welcome and Overview, Mickey Parish, PhD, Acting Director of the Senior Science Advisor Staff, CFSAN, US FDA, College Park, MD

If everyone could take their seats please, people get started. Good morning everyone. We are running a little late. I wanted to welcome you to the ninth in a series of Society of Toxicology, SOT FDA Colloquia on Emerging Toxicological Science Challenges in Food and Ingredient Safety. My name is Mickey Parish. I'm senior science advisor here at FDA CFSAN and director of the senior science advisor staff. It means that the longer the title, the less impact you have. Beside that, I'm here to try to welcome you and to say we are very happy to have you here today. This is a commitment that we've made with the Society of Toxicology. This is the third year of this colloquia series. There will be four colloquia in the year as we have done in the past. We are very pleased to have everyone here today.

I would like to start with a few opening remarks. I want to apologize for the fact that I am not a toxicologist. I am a toxicologist in training. Dr. Susie Fitzpatrick is making sure that I learn toxicology. I now know buzzwords like high-throughput and organs-on-a-chip. Thank you, Susie, for the education you have been providing. The colloquia that we have been engaging in for the last, now three years, are intended to provide a forum to engage in communications and discussions about the newest toxicological methodologies and to inform the toxicologists who attend the meeting by allowing them to interact with experts around the world. We do have speakers truly from around the world today.

The colloquia are not intended to offer a public forum to discuss or make recommendations about policy issues. But rather, they are intended to discuss the latest regulatory science related to toxicology.

The few administrative items for those who are here, the restrooms are down the hall to the right. There will be cards handed out for those who would like to submit questions for the panel that will occur at the end of the meeting. There may not be time to answer all the questions, but they will do their best to get through as many as possible. For now, it is my pleasure to introduce Peter Goering, the SOT past president, for his remarks.

Welcome from SOT and Introductions, Peter Goering, PhD, SOT President, US FDA, Silver Spring, MD

Thank you, Dr. Parish, for your comments and remarks and for helping us introduce colloquium. Behalf of the Society of Toxicology, we're grateful for the Center for Food Safety and Nutrition for being a partner, a collaborator in these colloquia. We are relishing in the success of these colloquia. Again, we have come together, our two organizations in a highly collaborative way to provide continuing education, high-quality, high-level training and continuing education for, not only staff here at FDA, but the impact beyond because of the technology and the WebEx. It has been a tremendous success. I think part of the success of these colloquia stems from the fact that the mission from our two organizations share a lot of similarities. The Society of Toxicology has a strategic priority to strengthen the relevance and impact of toxicology. On the other hand, the Food and Drug Administration have a strategic priority to transform toxicology to use the best approaches, tools to help support regulatory decision-making to improve public health. The Society of Toxicology has a strong educational and toxicologist development priority to develop and support toxicologists. The agency, the FDA, also has a strong emphasis on training its staff to be ready to confront the public health and regulatory challenges that confront the agency.

Here are the eight previous colloquia topics. The one at the top of the list was one that we held last May. The chronological order, starting with the first one on the partially hydrogenated oils, which is a major issue here at the Center for Food Safety. We've had a variety of topics, highly relevant topics in terms of risk assessment for carcinogens, issues and approaches like the threshold of toxicological concern, computational methods. All of these are highly relevant topics that are of high interest to the agency, high interest to toxicology and we are so pleased to be able to provide them here at venues like this.

This slide helps illustrate the success and the wide impact of these colloquia. This shows previous data for the previous eight colloquia. We have had almost 2,400 people participate by webcast, not only in the agency but around the world, in industry, academia, through the technology, people are watching this slide at 10 PM and 3 AM at different places around the globe. On site, we have averaged 60 individuals here in the auditorium. This slide shows the value added of taping these programs. These recordings are available on the Society of Toxicology website. You can see whenever we have a colloquium, there is a high increase in hits of people looking at not only the last presentation, the last colloquium, but the others that have preceded it. So we have up to, in some cases for the [inaudible] pharmacokinetic colloquium we had up to 4,500 people view that over the last couple of years.

We've had feedback every time one of these occurs. We ask for survey. We encourage all those present in the auditorium, all those participating by WebEx to fill out the survey to give us feedback. This help us to realize that we are on the right track with the topics and with the kind of presentations that are being made. So, we have made some very positive feedback about the quality, about the relevance. We appreciate the feedback.

Just some of the comments that people have made. People are very pleased with the presentations and the quality of the speakers and the knowledge that they bring. The presentations have been well done. They have been addressing training needs. Dr. Parish, please keep this series running as long as possible. People have been again impressed with the quality. It is a good professional development opportunity. A good continuing education opportunity, particularly in the time of tight budgets people are not able to travel to meetings as much as we were in the past. So it is a very good opportunity for that. Overall, just an excellent initiative.

By and large, the reason that the colloquia have been successful and of high interest is due to the SOT/FDA organizing committee. I would like to recognize this group. It is comprised of people both representing the Society of Toxicology and the people representing the Center for Food Safety and Nutrition. We owe them a debt of thanks for their efforts, chaired by Ivan Rusyn for a couple of years and then here, at the Center for Foods, Dr. Allen Rudman has been a real driver for the success of these.

Now, we will get started with the program. I would like to recognize Dr. Tim Shafer from US EPA who is going to chair the session today. Also, thanks to Dr. Jeff Yourick from the Center for Food Safety who has been instrumental in organizing this program.

Dr. Shafer, welcome to the FDA. We look forward to the session.

Speaker Introductions, Timothy Shafer, Chair, US EPA, Research Triangle Park, NC

Thank you, Peter, for the introduction. I would like to thank the organizers for the invitation to participate and organizing this, this particular colloquia and for inviting me to participate. I want to keep my remarks brief because we have what I hope you will find to be an excellent slate of speakers for this symposium this morning.

The topic of developmental neurotoxicity is one of particular importance. It is important to society. We have seen in recent years a dramatic increase in the number of neurodevelopmental diseases such as autism and ADHD. At the same time, we have thousands of thousands of compounds that we are exposed to in the environment for which we have little or no information on their potential to cause developmental neurotoxicity. We are concerned about these increases in neurodevelopmental disease and about the potential of the compounds that they are exposed to in the environment and the ideology of those diseases. This is one driving reason why this particular topic has been becoming more and more important in the public eye as well as in regulatory, with regulatory issues.

Dr. Chip Vorhees is going to provide an introduction to developmental neurotoxicity testing and also offer comments on some of the issues associated with the current paradigm we have. Then we will have a talk from Dr. Randy Peterson who will discuss the use of zebrafish as an alternative method for developmental neurotoxicity testing. There will be a break and then followed by two more presentations. One from Ellen Fritsche who owill talk about *in vitro* approaches to developmental neurotoxicity testing and screening. Dr. Anna Price will talk about the use of adverse outcome pathways and how they can be applied to regulatory issues associated with developmental neurotoxicity.

Without further ado, I would like to introduce our first speaker. Dr. Charles Vorhees is a professor in the department of pediatrics in the division of neurology at the Cincinnati Children's Hospital Medical Center at the University of Cincinnati. He is the former director of the graduate program in molecular and developmental biology and remains a director in their teratology program. Former director of their neuroscience program. His laboratory is focused on the developmental neurotoxicity of pharmaceuticals, drugs of abuse, and environmental agents, and he uses targeted genetic models in candidate genes related to depression and ADHD. He is a highly published author. He has more than 300 publications. He has served on a number of committees with the National Research Council, NIH, EPA and FDA. He's a former editor in chief of *Neurotoxicology and Teratology* and several different editorial boards. He received his Masters and PhD in neurobiology from Vanderbilt and did his postdoctoral research at the University of Cincinnati Children's Research Foundation before joining the facility there, or the faculty there, sorry. And he's co-director of the Animal Behavior Core at Cincinnati Children Research Foundation.

So, I'd like to introduce Chip and thank him for the time he's taken to be here today.

Developmental Neurotoxicity Testing: An Introduction to the State of the Science and Opportunities for Improvement
Charles V. Vorhees, Cincinnati Children's Hospital Research Foundation, Cincinnati, OH

Thank you, Tim. It's a pleasure to be here. You can hear me. I think there's a little bit of feedback. It's a pleasure to be here to talk to you about DNT, developmental neurotoxicity. Technical challenges.

I'm going to jump right into my, first I have no conflict of interest in the presentation. I want to remind you about the complexity of brain development. Some of the other speakers will go into some of the complexities of brain development. I want to remind you of the timeline of brain development in relationship to the development of other organ systems. You can see it is very extended as opposed to the development of other organs that become issues in teratology. This is an old slide, [inaudible] Corporation in two different brain regions that was published many years ago. It shows you relative neurogenesis in different brain regions. In rodents continue postnatally even though those are prenatal events in human embryogenesis.

Things that are not covered by neurogenesis, I adapted from a review, trying to show you neurotransmitters have a developmental profile that is very distinctive and here I use this translatingtime.net site to try to give you an approximate comparison. There is no exact comparison between stages of brain development in utero in humans and rodents. So, you get to these problems of when do you expose things and what is that you are mimicking when you try to model exposure?

So, I just try to give you some human equivalents. This is post conception 8 weeks in humans. It's about embryonic day 15 in a rat or mouse, 13 translates to about E19. But then when you get into third trimester, late second and all throughout the third trimester, where is the equivalent to birth, if you will, in humans in a rodent? Some people called postnatal day 10, some call it day 12. Some say it's later than that. So, it is actually, there is no equivalency point. So, this is something that you have to constantly be aware of. So when you read a paper and somebody says we exposed out to postnatal day 10 because that's the equivalent of birth, you have to say to yourself, yes and no.

The complexity of brain development creates assessment challenges. Assessment challenges at the DNT could not possibly cover and was not really intended to entirely be able to model, at least as I look at it. So, has the complexity in brain development led to, been translated to the design of the DNT studies? The answer is, it has not, and it could not possible in a sense because the brain complexity is so much greater than what one can do in a DNT study. And, when you think about when the DNT study was introduced, it was introduced long before many of the complexities of CNS brain development were even understood, as if they are today but they are much better.

So, I am going to give you some opinions here as I go through this. One of the things I will say about the DNT is that it is based on rather simplistic observations, many aspects of it. It tries to cover basic movement in the open field, spontaneous exploration, habituation. It is a little more sophisticated when it looks at the acoustic startle reflex because it is a complex reflex. It is often called ASR for short. When it comes to learning and memory, which are supposed to be its assessment of higher functions, it really is still very rudimentary. I will get to why it is rudimentary in terms of what kinds of learning it assesses and what kinds of memory it has evaluated.

Why is it rudimentary in this way? First of all, when these guidelines were written, there were no methods to tackle the complexity of CNS development. We will hear later talks there are approaches come down the line that began to get at these greater complexities. Because guidelines were general, described as being flexible to allow people doing regulatory DNT studies to use their scientific judgment to do the best possible job. The problem when you make things really flexible as you run the risk of getting the lowest common denominator kind of assessment. I actually think this flexibility which was put forward to be a strength, I think is a liability. I think it allows for too much flexibility.

So, does it have to be the way it is? I would say it does not. Here is an example taken from paper by [inaudible] at EPA. It's a typical design of an EPA type developmental neurotox study. You can see the various assessments. I put the abbreviations here, but I think you already know them. Functional observation, battery open field, learning and memory, acoustic startle response.

The very first thing that you notice when you look at this FOB, this DNT, and you look at the FOB part of it, you have to ask yourself, who picked these days and why? What is the biological basis for assessment at four, 11, 21, 35, 45 and 60? What is this based on? Is there a scientific literature basis that supports these? I can't find it. The open field. Why are the assessments at 13, 17, 21 and 60? 13 is very problematic. Rats don’t really move around a whole lot at postnatal day 13. You get very low numbers. Is there really a difference in the animal’s exploratory habituation patterns at 17 versus 21? Not much. So, the question is, are you really getting uniquely useful data out of doing it four days apart? I'm skeptical. Why is there acoustic startle at postnatal day 21 and 60? I've done a lot of acoustic startle testing in my lab and basic research settings. This age is a little difficult. You can get a startle response, no problem. But it is quite variable. Why would you do a test at an age when you know the data are going to be highly variable when you can get very stable data if you test somewhat later?

When it comes to the learning and memory test, this one is logical, but why this age? Why postnatal day 21? The learning repertoire of rats at postnatal day 21 is extremely limited. You can only do very rudimentary cognitive assessments at that age. It seems to me, if one of the

goals of the learning and memory aspect is to get a higher cognitive function, which is a very much a human characteristic that you would want to protect for, then why do something that's inherently limited at this particular test age? I went over these and I won't go over them again. These are the common tests that you see at least in the EPA DNT studies which I will come to, the ones in pesticides at least. These two tests actually have a significant number of problems. I will show you for passive avoidance why I think these two tests are inherently very limited in their sensitivity.

First, I want to talk about the FOB. I took these two FOB's out of two different papers and I reference them in the reference list I've provided. Over here, you could see this laboratory's list of FOB parameters. Over here is another laboratory's, also called the Erwin laboratory. What was the idea behind it when Erwin developed it? It was supposed to be a quick, simple screen for drugs. I will ask you, if this looks quick and simple? I don't think it meets those criteria that he originally set forth in his early papers. What I've done here is circled in red all the tests between these two laboratories that are convergent. They do the same test, look at the same parameters. It looks pretty good, right? Here with that are divergent. Over here, they are not over here. Altogether, there are 31 nonconcordant items in these two FOB's. I ask you, is the FOB a thing? Is it a entity?

Here's a little quiz. Based on what you just saw, what is the FOB? Can you tell me back what the FOB is? You cannot because the FOB is not a thing. It varies from laboratory to laboratory and what people believe it to be and what the measure and how they measure. Is there any central processes that underlie these tests? What are the central nervous system constructs that are being evaluated by using the FOB? I do not see any. What kind of data do you get? Here is a paper I took to show you some FOB results. For each of these tests, the animals were scored and their average rating which is a subjective rating of severity of each symptom is shown. These are for pesticides.

Next quiz: what were the findings? Can you summarize what was found in that study? I do not think so. It's hard to know what was found. What were the problems with the FOB? It's subjective. Not truly quantitative. You could argue that the scales are attempting to be ordinal. Data analyses, what you saw the table typically is what is done. There is no statistical analysis done of those kinds of data. There have been attempts made to convert them all to Z scores and put them together in clusters and do statistics on them, but the papers I read that attempted to do that has generally come up with, well we didn't really find anything.

So, the items that tend to be unrelated, they are time-consuming, there's no underlying construct. I will come back to the issue of how sensitive is the FOB in terms of reliability and validity, there really are virtually are no studies on this issue. There are few interlaboratory studies that have been done to see if there they are valid. The reliability issues which also goes across laboratories. If I test this particular pesticide, drug, or food additive over here and it's tested by another laboratory over there, and they both use the FOB, did they get the same finding?

The first problem is: can you even find such cases in the literature? The general answer is no. I just want to run through motor activity open field. This is just our setup. This is what it looks like. If you have not seen it, this is an example data. This is the exploratory phase going down. Eventually you get habituation. This is with a early exposure to citalopram in study we were doing. After you give them, this is 30 minutes with no treatment. They were exposed early in development and they are now adults. Over here we gave them an injection of saline and in the third phase we gave them an injection of a drug challenge to see if we can unmask latent

effects, I'm not going to show that. So, the sensitivity is pretty high. It's one of the tests that often shows an effect. What that effect means is a regulatory interpretation issue. Its reliability is moderate. Its validity is also in the moderate range. Its type I error rate I would say is moderate. We know that open fields produce type I errors. They produce false positives. We know that. You often see it in studies where it is done once, and you get a low dose effect. You do it again and you don't get a low-dose effect, maybe you get a mid-dose effect. We know there's something going on there. Or you get no effect next time. It is apical, so it is a very nonspecific effect and the advantage of that is that it suffers a lot of underlying functions, but it is not specific.

I took this out of the old FDA collaborative behavioral teratology project. These animals were exposed, this paper so old that you cannot get a high-quality reproduction of it out of the Journal. They were exposed to methylmercury during development. This is just to show you that the laboratories get surprisingly consistent results. Actually, all the laboratories detected the higher dose of methylmercury. The low dose some laboratories detected as significant, some didn't.

This is to show you acoustic startle. That is a typical apparatus. It is not a plug for that company at all. This is a typical example of what kind of data you get. This is the peak response. This happens to be a drug study. It does not matter with the drug is here. This is a significant effect. You can see what that looks like when you get suppressed startle. It has a lot of aspects of validity, of construct and predictability. It has been used a lot and humans as well as in animals. The nice thing about it, you can test the same reflex in both humans and animals, so you get nice cross species comparability. It has content and face validity and empirically, it has been used, and I will come back to this point in the Office of Pesticide Programs studies which I am going to talk about here in a moment, it served as the part of departure 19% of the time.

The line does not show up here for some reason. This is pre-pulse inhibition. What you see is the basic startle response here. This is the knockout of [inaudible] a gene we were studying in a knockout model. And then down here, this is what happens when you give a weak pre-pulse that precedes the strong stimulus by a certain number of milliseconds. This is the pre-pulse intensity. The numbers are blurry, I'm sorry about that. You can see that that suppresses the response. That is an aspect of attention, or often called sensorimotor gating. It's used a lot in human studies and animal models of various neuropsychiatric disorders. It's quite a useful measure and you can get it by running basic startle on then by running this. It would add 20 minutes. So, you are doing this huge DNT study, and for 20 minutes' additional testing, you could get this kind of data.

This is basic startle from the FDA sponsored collaborative behavioral teratology study to show you what startle habituation was found in these different laboratories. Again, all laboratories found the effect of high-dose, most laboratories found the effect of low-dose.

I had a hard time finding an example of a water T-maze is which is widely by CROs. They all look different. They are different shapes. There are T's and ends. Basically, it is a two choice. What I would like you to notice is how close the data are from the very beginning. Yes, they got significant effect from the compound they were giving which is a compound you would expect would produce effects. My point is, look how rapidly this task is learned. By the fourth trial, even in a compound that is effective, they have caught up. I would argue that test like this are not particularly optimal if you are not trying to get higher cognitive function.

So, this is an apparatus that shows passive avoidance, and again this is a paper that I took off the Internet. I'm trying to show you the kind of variability that passive avoidance produces. This is an issue because the more variable your test, obviously less sensitive. Here's an example from my own lab. We do development studies all the time. This is a study on the developmental effects of methamphetamines, animals that were exposed early in development, then they had no exposure until they were adults. They were put through a series of learning and memory tests. Here is the Cincinnati water maze. You can see all the animals that were treated were impaired in their ability to learn this task. This is run in complete darkness, under infrared light, they cannot see anything. They have to use internal navigational cues to get from the start to the goal, it takes some a lot of days to even begin to learn it. And actually, we now test it out to 18 days. This is a Morris maze. It's a picture I took of the Internet of it. The concept is you have to use distal cues. The first one is internal cues. This one relies on distal cues, cues outside the equipment. You start them from a different location in the goal, the hidden platform which is submerged and camouflaged, is where they have to go. Because they are always starting in a different position, they cannot memorize a route. They have to navigate from knowing where they are in the room. You can see the developmental methamphetamine uniformly impaired their performance on this task.

I did not show you the radial water maze. Originally, it was a food motivated task. It can be adapted to be a swimming test. In that study, we got significant effects, in the Cincinnati water maze of egocentric learning, in the Morris water maze of spatial learning, in the radial water maze of working memory, and what is passive avoidance in the same animals.? Can any means be more similar than those two? Nothing.

Trying to find studies where people have actually compared passive avoidance to other learning and memory tests are very hard to find. This is one reason I did that study. People kept saying where are the data, where are the data showing that passive avoidance is not a very sensitive test? And finally, I said, OK, I guess I’m going to have to do it. That is the first data we have that kind. We are testing passive avoidance against those other tests in some other studies, too, that are ongoing.

This is a review at EPA after they got through a lot of DNT studies for pesticides submitted. They did a review. They asked the question about DNT. They had all the other toxicology tests and so they asked the question, what does the DNT test add to what they knew about determining the point of departure for the risk assessment? So, locomotor activity open field, they call motor activity, provided a point of departure 29% of the time. That's useful. Acoustic startle, 19% of time. The FOB, 4% of time. I would argue that when you're talking about pesticides which are neurotoxins, obviously, to begin with, that's not so good. Learning and memory were also about the same. It was also very low. However, I would argue that the reason those two tests performed poorly are different.

Am I saying that the FOB and learning and memory tests are insensitive, they will not pick up pesticide, what use are they, should we just throw them out? No, I'm not saying that. Not necessarily. I'm saying that the problems in the two tests types are different. One of the problems with FOB is it’s not one test. I've shown you that. It has many moving parts. Those moving parts are not even the same from one laboratory to the next. So, it's not one thing. It's very difficult to compare data. I think it is in striking need of revision, simplification, focus on those tests within it which can be shown to be useful and are quantitative, and a lot of the really subjective tests of which we do not know anything about the CNS substrates, should probably be eliminated. This test made what it was intended to be a quick screen.

Learning and memory. The problems with the tests are the tests that were chosen. Water maze test, passive avoidance, and we choose better tests we make a better data.

Why does this test constantly, why do people always discuss test and its problem? I think there are reasons. This slide I got from Tim. I appreciated very much. This is the problem that are faced by chemicals in commerce. At least 10,000 chemicals that have to be [inaudible] at least by the Environmental Protection Agency. How many of them have developmental neurotoxicity data for? Less than 1%. The question is: could you ever catch up? The answer is no. The fastest you could do DNT, there are more chemicals entering commerce. The scale of the problem for the pesticide program are different. As far as I can tell there are hundreds of pesticides. They are tested at least in that EPA study, 79 chemicals cause 69 of which they use. Is it possible to do DNT sent every pesticide ex-cop apparently. The data would suggest it is feasible.

They have the regulatory authority to do so. That is the issue the same at FDA? I think it is very different. So, to try to put this into context for this agency. I went to the FDA's website. I looked at how many food additives the agency regulates. According to the website, it's 2000 direct, secondary direct, food additives, GRAS and prior sanctioned substances which I assumed were things that were grandfathered in. His DNT manageable for that never? It probably is, especially if you did them gradually. Perhaps you have data on them, I don't know. It is not an insurmountable number. Depending on how many you get per year, that's an interesting question that I don't have data on.

At your related center on drugs, about 1500 since the FDA was founded in what, 1938? That's what your websites has. You regulate 1500 pharmaceuticals and you get 30 per year. That's very manageable for DNT. They are putting them into people that have a biological effect and you naturally would worry if they produce adverse effects on brain development. When it comes to the residues problem, yes. You have the same problem the EPA does. It is a huge amount of residue that potentially appears in food for packaging. Their DNT is definitely are a problem. In conclusion? The problem as I think DNT's across agencies and programs varies widely. You can do them routinely, others you cannot. You have need high screens just to cope with the sheer volume.

At FDA, I interpretation of what I see in the data for the drugs, they are feasible to do on all drugs, really. You may want to prioritize them. It may not be a cancer drug, I don't know. For food additives, yes. If some priorities are set. [inaudible] that will allow you to set priorities and select only those that came through-put through screen. At EPA for pesticides, you reasonably can do every pesticide that has ever been marketed. For chemicals in commerce, clearly not. That is where I think particularly the need for high throughput screens and AOPs come through. Just as we do gene screening for some things, you can screen now for tens of thousands of screens, but when you get your heads, you still have to do RT PCR's to confirm. I think that's what you have to do here. You use the screening methods, you find the suspicious players. Then you come back and do the DNT for confirmation to give you a sense I'm not going to go into the issue if they don't agree. That's a complex issue.

The need. This is another slide from Tim that was taken from this review paper. And I thank Tim. I have read others but not this one. It's interesting. They show some data that in several countries there is a clear increase over time in autism, attention deficit, Tourette's. In the states, we know increase in autism spectrum disorder there's been a pattern of the listed delegates. You can go into a debate as to why that is. It is clear there is a widening definition

of the disorder. Some of it is that people are paying more attention and watching for. They're been interesting studies that show that as they the reported prevalence of ASD has gone up, the reported prevalence of learning disability diagnosis has gone down almost as a converse match. There is some question about what the change in diagnosis and perception of ASD versus learning disability impacting what we think what we used to call learning disabilities. My view is that the FOB is inconsistent and content. It is very objective and insensitive. Activity is good. Acoustic startle habituation is good, but you can ask add pre-post inhibition to the whole procedure perk learning and memory needs and overhaul and I think switching to better tests that's the way to do that. What makes for good learning test? One of the problems I've just shown you are the ones that have been the most prevalent use are too easy. That has a steep learning curve which makes it insensitive. Flat learning curves and nobody is learning anything

You don't want it to be aversive. Shock tests are aversive. You want to make the motivation something easy for the laboratory to implement. You want to be able to equalize the motivation of cross groups. Many compounds given the developmental effect, growth, and body weight. The animals, when you're testing them you use body weight and food reinforcement, how do you equate for appetite and palliative itty. Fundamentally, you cannot. So, a solution I'm going to suggest is swimming tasks you can choose any test that you use for learning and memory you must be able to separate out performance from learning factors. You need to look at relative types. This is just the general textbook example to the types of learning we are going to focus on instrumental learning. What types of memory, memory has been [inaudible] explicit memory, declarative memory, subdivided we won't get into right now, and implement all memory for skills and has.

Driving to work, makes you pay attention. You can almost do it in your sleep, I would not recommend closing your eyes. You can get to work and it's a highly rehearsed skill but it's a complex skill. We can assess that and animals with the right test. So, for declarative, this is also studies in animals called allocentric. In people, it's people for people, places, things and events. It's oriented to surroundings and has distal cues. The Nobel Prize was given to the studies in these areas. We know it included hippocampus and entorhinal cortex. We know a lot about what mediates this type of learning and memory. For egocentric it's been studied less. Its memories for skills, paths, events, order. Self-orientation, you use proximal or internal cues. We know the brain region involved is the striatum. It's also presubiculum and others. It overlaps with the hippocampus and entorhinal cortex. But that is how the brain does it. It uses redundant networks. For allocentric: the most common methods are the Morris, radial-ram, radial water maze. First burned to have to put a lot of time and effort to make this work. You just can't set it up and run them. Egocentric: we use this transfatty water maze that you can use the cues Morris maze. There's foraging maze. The Wishaw method is a variation of something like the Barnes.

I would suggest that water mazes have significant advantages and regulatory studies. They do not require trainings as rodents are natural swimmers. Humans have to be taught how to swim. They can be slow when it comes to survival skills. On the other hand, animals will survive well. There's no unknown species you can put in that they will not just start swimming by instinct. In the water maze, you have to do a minimal amount of training. They are immune to appetite differences. Obviously, within a wide range they’re immune to body weight differences which is sent advantage. The motivation is intrinsic. And 100% of animals will do the task and competitive maces you have a dropout rate. [inaudible] what you make out of that? In these mazes, you don't have that problem. In nearly 100% of the animals eventually master it. If you run enough test, virtually all of them learned.

This is an old one, this is a new one. If you haven't seen them in person. We've gone to a larger when the most vividly is. Most people use 6 feet in diameter. So about 183 cm. We've now gone to 204 cm, a peeping diameter. Like you think, if you open the space you make it more difficult and sensitive. This is the Cincinnati, the fisheye effect distorts the look of it. Just to give you a sense of what it is like.

I think there are two types of improvements needed. I think the tests need revising. I think the test ages need revising. There are too many test ages and many of them are not very well founded in central nervous system development. They need to be designed to fit the agency's needs. I think having one FOB is suggested for everybody is not the past. They are not mechanistic, and they are not intended to be. We will care about mechanistic studies. Really, DNTs were not designed to be mechanistic. They are designed to assess outcome or impact. If we pick tests better, they can provide us hints. I show you this is a nice review paper that shows the acoustic startle reflex down to millisecond delays as the reflex comes into the auditory system and travels up to the nucleotide in the brains that brainstem. Is very carefully mapped and we can use that to give you some mechanistic information if you care to do so.

The Cincinnati water maze these are some studies we've had showing that it is [inaudible] dependent producing [inaudible] in various subregions of the striatum that produce striking effects on Cincinnati water maze performance. This is from the Nobel work of the Moshers showing how the very complex grid cells that mediate spatial navigation. If you use those tests, it would give you information. At least you'd know what part of the brain to look for. histological or genetic gene expression changes.

I think for future direction, it is not appropriate in all contexts. It is often necessary but seldom if sufficient. It is not mechanistic, but it is sometimes the only option you have. And until some of these other tests become accepted that you're going hear about next, you still have to depend on it. But it needs to be revised. I think it should include pre-pulse inhibition as well as startle habituation, Morris water maze. Should other studies designed include these kinds of assessments? There's a proposal out there, I think it is from OECD, extended one generation reproductive study that would capture a lot of the things and illuminate all of these individual studies. They are proposing cutting learning and memory out. Really? What you really want to do that? What is the target species? Humans. What are humans known for? Thinking, reasoning, attention, learning, memory. higher function. Why would you cut the higher functions out? I don't know. That baffles me.

This is my moral slide. Brain development is complex, and many things can go wrong, but the effects in humans when people do studies in humans for toxic effects, what did he end up finding? Increased forgetting, poor recall, import learning content. Problem-solving, attention problems. All cognitive. So, the DNT is weak and cognitive assessment. If the target species is human and these are the kinds of things that people find type think about and developmental lead, may be desired to extreme examples, but what are showing up in these developmental studies? These kinds of things! The way you get at these things is by assessing comparable higher functions in animals and PPI which gives you an aspect of attention. It's one of the major areas that goes wrong. These assessed two forms, we have to stop this idea that when we say you have to test learning and memory. Learning and memory is not one thing. It's a whole complex set of things. People have short-term, long-term, working memory is a short-term kind of memory but in terms of long-term you have different kinds. You have allocentric space, declarative, and you have the learning of skills and locations.

This is my bottom line. There are 69 regulatory studies done on pesticides. These studies are telling you something. And are we listening to what that review says to us? Or are we going to sit around and do 69 more studies the same way? After 69 more will we finally say, we have some tests in there that are not giving us what we want. When is enough enough? To change your thinking and start doing things differently? I think we are there.

References. I want to thank my longtime collaborator Dr. Michael Williams. I want to thank NIH for supporting my research for all these many years. Thank you very much.

[Break in audio]

Shafer: research Center and a senior associate member of the of the Broad Institute. Dr. Peterson's lab specializes in using high-throughput screening technology with zebrafish to screen drug candidates for cardiovascular and nervous system disorders. He's going to talk with us today about his experience using zebrafish to screen chemicals for effects on the nervous system and cardiovascular systems.

Zebrafish as an Alternative Species for Developmental Neurotoxicity Testing that can Provide Hazard Identification and Mechanistic Information
Randall T. Peterson, Harvard University, Boston, MA

Thank you, Tim. Thanks to all of you for participating in this symposium. Is that better? It's a real pleasure to participate. I appreciate the invitation. I'm happy to talk to a little bit about how zebrafish as an organism is contributing or potentially could contribute in the future to some of the challenges that we are talking about today. Like Dr. Parish, I am an enthusiastic learner of toxicology, not a card-carrying member. Much of what I talk about today it's discovery. I hope you can use imagination to see how some of the approaches and technologies that have been developed in that area could be applied to toxicology.

As disclosures, I want to disclose that I've received funding and equity from a variety of companies and funding sources. Most relevant for this talk is that some of the research I will be telling you about was funded by Hoffmann-La Roche and the National Institutes of Health. And some of it was done in collaboration with Roche and with Teleos Therapeutics.

Fundamentally, what we are all trying to do is find a way to define the activity of chemical compounds. Of course, if you're interested in one simple target or biological phenomenon that can be studied *in vitro* you can see generate these beautiful dose response curves that describe the activity of the compound in that assay. But, as we all know, real life is much more complicated. Particularly when you start thinking about something as complex as the brain where you have multiple targets, many unknown targets, and complex interconnected activities. It just becomes much more difficult to describe in a meaningful and complete way with the activity of the compound is on this complex system. So, we like many other others have been interested in trying to find better ways to define the activity of compounds I complex systems like the brain and, also more efficient ways to identify compounds that might have detrimental effects on the nervous system.

Really, many toxicologists and other researchers have been faced with the difficult choice between complexity and throughput. You can imagine choosing an organism like a rodent or a

human to study things in that offer the benefits of having high physiological complexity. But of course, low throughput. Others have tried to choose a more *in vitro* system where they can have much higher throughput in their studies, but it the expense, often, a greatly reduced physiologic complicity. One of the attractive things about the zebrafish that it offers some of the best features of both worlds. You have a complex physiological system with an active brain, complex brain, and yet, you're able to do experiments and the throughput and scale that you can do with other *in vitro* systems.

Let me just tell you a little about the zebrafish as a model. What's attractive about it is that it enables you to do *in vivo* biology and at the scale and cost of the *in vitro* systems. Many institutions have these facilities that enable us to house tens of thousands of adult zebrafish inexpensively. Animals are highly [inaudible] so each female can generate 200 fertilized eggs a day. So, from a facility like this you can routinely generate tens or hundreds of thousands of new offspring a day. They are very small. What I'm showing you here is a single well of a standard 384 well compound screening plate. You can see a couple of zebrafish embryos developing in one of those wells. They can survive just with water and the yolk sac that they possess for days.

Another great advantage of the zebrafish is that over the last couple of decades, a very rich repertoire of phenotypes has been described. Literally thousands of genetic perturbations of distinctive phenotypes. But that enables is if you screen a compound and you discover a new phenotype that is produced, you can search the databases and rather quickly identify hypotheses about perhaps the pathways that have been disrupted by that compound. That is simply by comparing the phenotypes.

Another nice advantage of the zebrafish is that they are highly transparent. That enables you to see what is happening both morphologically and physiologically inside the animal without having to cut the animal open. I have highlighted a few examples in one area. This is the neuronal vasculature. You can also see that there are transgenic lines that enable you to highlight the vasculature of the brain and then very easily identify perturbations in the normal vasculature. Because of all these advantages, zebrafish are increasingly used in toxicology for echo toxicology, developmental toxicology, cardiotoxicity, various other organ specific toxicities and increasingly in the nervous system as well. As a result, the number of publications using zebrafish has really exploded over the last decade or so. They are now almost 30,000 publications that reference zebrafish.

I know this audience is particularly interested in the safety and efficacy. It turns out there's actually a quite later check in this area as well. The use of zebrafish for these studies, most of it has emerged in the last five or 10 years. This is categorized into different ways. On our studies of food contaminants including things like pesticides, endocrine disruptors, natural products, various toxins and whatnot. Nano materials and byproducts of food production such as acrylamide. These are studied in the zebrafish. Another group of people who are studying the positive effects of foods, and things like antioxidants, neuroprotectants, [inaudible] and metabolic modulators have all been studied using zebrafish.

If you're interested in this area, I would point you to the search for you which was published recently. I think it is a fairly comprehensive of the application of zebrafish and to these types of studies. I won't go into any of them in great detail here. Almost all of these studies study in detail on particular compound or substance or a small group of them. I think those studies really highlight the level of detail and relevance that you can generate with these individual studies. My own personal interests in what I have gone toward high throughput studies in one

of the real advantages of zebrafish is they enable you to study these things in much larger scale. So, for the rest of the talk, I will focus on high throughput application of the zebrafish and how you can study the effects of compounds in rather higher scale.

The very first example that we tackled in our group, this is in collaboration with David Milan, Calum MacRae, and Goeff Burns, was an investigation of high throughput cardiotoxicity. The approach that we took was to engineer zebrafish so that their hearts would fluoresce green. That was done by expressing GFP from a cardiac specific promoter. You can see here, this is you cannot see the whole embryo, but here's the zebrafish embryo, the glowing heart here. We developed automated microscopes that could take 96 plate filled with zebrafish and progress from one well to the next to identify the fish and the well based on the fluorescence of the heart, after a short video of the heart beating, and then some week to [inaudible] transfer of the life intensity in a very precise heart rate measurement for each fish in each well in this 96 well plate.

With that technology, we were able to test the effects of 100 on drugs on the heart rate. We selected 100 drugs amongst which we included 23 that were known to have a liability in humans for prolonging the QT interval. In other words, cardio toxic compounds. We asked in an automated way whether this system could identify the cardio toxic compounds. I realize the text here is too small here to read. Each of these red compounds was a compound that was known in human to be cardio toxic.

What we found is that the system is able to identify 22 out of the 23 QT prolonging drugs in an automated way in the system. That really gave us confidence in two things. First, we could take even the complex physiologic phenomenon like QT prolongation and reduce it to a high-throughput automated zebrafish assay. The other thing that it really gave us confidence was that there could be a very high degree of conservation between the effect in zebrafish and in humans. In this case, we were able to identify over 95% of the QT prolonging drugs.

That emboldened us to look at other phenomenon that might be of interest to study in high throughput. One of those is neuronal toxicity. We wanted to develop a system where we could, and a very high throughput facile way identify [inaudible] cells in the brain of zebrafish. We decided to exploit a phenomenon that was well where in the plasma membrane of a cell, there is an asymmetry. The phosphatidyl serine is localized in a living cell exclusively to the intracellular leaflet of the plasma membrane. What is known is that as the cell begins to become stressed and undergo apoptosis, the phosphatidyl, the asymmetry in the phosphatidyl serine is lost. And you see it appear on the outer leaflet of the cell.

So, we were able to generate an annexin V fusion to YFP. We then generated transgenic zebrafish where this fusion of secretion signal annexin V and YFP would be expressed ubiquitously in the animal.

You can see here is the head of the early embryo and the body goes back like this. We found that cells begin to apoptosis would collect the secreted protein and it would give us a very bright and clear indication of which cells were under growing a proptosis. Unfortunately, we were unable to get videos to work on this AV system today. This was supposed to be a video showing you development of the embryo. You can see from the first frame that most of the embryo is dark. There are few cells that are beginning to undergo a proptosis. You can watch in real time as new [inaudible] cells emerge and, in some cases, migrate and some cases are engulfed or lysed in the animal. I'm showing you a genetic perturbation. The same is true with a chemical perturbation. You can identify regions of the brain that are collecting [inaudible]

cells. Here we're looking at a confocal slice through the forebrain of the animal. Even in the healthy animal there are a few [inaudible] cells, that's normal. But with a perturbation you can see abnormal accumulations of [inaudible] cells.

More recently we are trying to use behavior as perhaps the ultimate readout of changes in neuronal function. We have heard a beautiful talk this morning about how behavior studies have been a great way of identifying neurotoxicities. The zebrafish are nice for this as well because they exhibit a rich repertoire of behaviors. These include simple things like stimulus response behaviors, but more complex things like thing sleep, learning, fear, aggression, social interaction. All of these things can be studied in zebrafish. Although there are some subtle differences in the way that these behaviors present between fish and humans, the gene cells and circuits that enable these behaviors are similar to those involved in human disorders.

We wanted to try to develop the capacity and a high throughput way to measure interesting behaviors in zebrafish. This work was really pioneered in my lab by Dave Kokel, who is now a professor at UC San Francisco. And more recently Andrew Rennekamp, who is currently a postdoc in my lab and looking for a job right now. They had to develop new tools, assays and most importantly analytical tools for recording these behaviors and identifying perturbations. I will walk you through a little bit of that now.

One of the first tools built by David Kokel was something that he called sauron, after the all-seeing eyes in the Lord of the Rings movies. Essentially what it is a tower with an eye at the top. A camera looking down some nice optics at a 96 well plate. And in this 96 well plate, we put 10 free swimming zebrafish larvae in each well. There is infrared light that is transmitted up through the plate to the camera. Then there is the stage here which can deliver a whole series of different stimuli. These can be different wavelengths of light, sound, other vibrations. Even electrical shocks can be delivered in a precise way. Then you can record through the computer the responses of the animals to those different stimuli. So, Dave developed 10 different behavioral assays. This sauron instrument is able to in one day look at 20 plates, 96 well plates of animals, being assessed through that battery of 10 behavioral assays. That can allow you to look at 2000 compounds or unique conditions. With 10 animals per well, that's 20,000 fish assays per instrument per day. This is now a routine with multiple of these instruments going for us to screen 20,000 to 60,000 animals per day.

I wanted to show you a video to give you a sense of the quality of this. Again, unfortunately the videos are not playing. But you can get a sense of what a 96 well plate looks like. This is a standard size drug screening plate. These have 8 to 10 drug free-swimming animals in each well. What you would've seen had the video function, the animals would be slowly swimming around, minding their own business. When a stimulus is presented to them, they all in unison undergo a characteristic response to that stimulus. Of course, all the while, the movements of the animals are being tracked and very rapidly, almost in real time, the computer is then able to extract the relevant data and plot the movements of those animals in response to the stimuli. We are able to identify different wells where there is a perturbation in the way they responded to a particular stimulus.

Even with a very simple stimulus, such as an acoustic startle, you can really measure multiple things in a rather rich way. An example is the faint blue bars. They represent acoustic stimuli that are presented. The black lines show you the movement of the animals over time. You can see that the zebrafish respond robustly to the three acoustic stimuli. We then present three quiet stimuli, sub threshold similar. You can see that the animals do not respond to those sub threshold stimuli. Then we applied the loud stimulus again. This time, 80 times in a row in rapid

succession. What you see is that they respond at first, but they quickly habituate, a form of non-associative learning, habituate to that stimulus and stop responding.

Then you can use his high throughput format to screen through thousands and thousands of compounds and identify those that perturb those various features. You can find compounds like this one shown here, which changes the duration of the response to the loud stimulus. Or you can find compounds like this one shown here which actually sensitize the fish and enable them to respond to that sub threshold, quiet stimulus.

You can find compounds like this one here, which prevents the habituation to the repeated stimuli. Depending on what you are particularly interested in how you can use this in the drugs discovery mode to discover specific compounds with specific features. We think perhaps the most exciting applications of this kind of technology come from creating a behavioral database that integrates the activity from lots of different kinds of assays. You can imagine taking multiple essays, screening multiple compounds, thousands of compounds in each one, collecting them in the database and then giving them a rich description of the activity of the compounds in this whole battery of different assays over time. Then if you had a suite of analytical tools that could help you identify compounds with desirable compounds, if you're interested in drug discovery, or undesirable compounds, if you're interested in safety assessment.

Just to give you a sense of what that looks like in real time and a specific example I'm showing you a concatenation of 10 different assays that will run in succession in the sauron instrument. What you should be able to see in a gray line is the movement of animals across these 10 different assays. This is about 20 minutes of laboratory time. These different colored bars at the bottom show you the different kind of stimuli that are being presented. And on the y-axis, you see movement. And so, in gray, these are DMSO or control-treated animals. You can see that they respond to different stimuli in different ways. I have overlaid on that haloperidol, an antipsychotic. You may be able to appreciate that for many of the assays, haloperidol does not change behavior. There are certain assays we get a significant change in behavior. So, we think that for each compound your essentially create a fingerprint that represents the complex effects of that compound across his battery of behaviors. I can tell you that the signature for haloperidol is unique to haloperidol you can take it out of a list of blinded compounds simply on the basis of the signature that it produces. It also enables you to create a large library of signatures representing thousands or tens of thousands of compounds with known and unknown pharmacology.

Then, if you have a compound of particular interest, you can run it through the battery of assays, create a fingerprint or signature and then use the computer to search the database, much as you would do if you have a genetic sequence that you wanted to search the genome. Here, you can take a behavioral signature and search the collection of phenotypes and identify other compounds or genetic perturbations which produce a similar phenotype.

We think this is a powerful tool for drug discoveries, but also for identifying mechanisms of action of a toxic compound. Let me give you a couple of examples of how this works. Here is an example using MK801 which is in an NMDA receptor antagonist. We took this compound and tested it at a variety of different doses in high throughput, ranging from 10 micromolar to 640 micromolar. And we got a signature for each of these different doses for MK801. And then we asked the computer to search the database of other compounds to try and find things that look similar to it. Again, we call this “PhenoBlast” or it is similar to a genetic sequence blast that you might do. This heat map here shows you areas where we got strong matches to other

compounds. And what perhaps you can appreciate here is that there is a strong, there are really three strong matches that came up in the search. Two of them are other exemplars of MK801 that already existed in the database. The third match is [inaudible]. You may recognize that is a different class of NMDA receptor antagonist. This suggests that you can find identical compounds, but you can find it structurally divergent compounds that work through the same mechanism.

Here is another example using duloxetine, again at doses ranging from 0.6 micromolar to 40 micromolar. The computer was able to find to several other matches and drugs. Interestingly, duloxetine found other exemplars of duloxetine, but also found fluoxetine, a structurally similar antidepressant, and compounds like [inaudible] which are structurally divergent antidepressants that do similar things functionally.

Here's another example in the antipsychotic space, and again, I apologize that these are drug examples, not toxic compound examples. But I think you would see very much the same thing there. Clozapine at different doses, you search the different database and you find matches to several different compounds. The strongest being [inaudible] which is reassuring. Also, olanzapine, which is structurally very close. That makes sense. Interestingly, haloperidol and risperidone, structurally divergent antipsychotics, also appear as weaker matches in this way.

That gave us a sense that we could take known compounds and understand the mechanisms. It also encourages us that we might use this in drug discovery. Here is an example where we took haloperidol, an antipsychotic, we generated a signature for it and blasted that against 30,000 compounds in our collection. We took the compounds that seemed most similar to haloperidol. There were 47 of them. We looked at what those were, they included five known antipsychotics which was reassuring. And 33 novel compounds. When we looked at those novel compounds, six of them were structurally related to each other. I'm showing you the structures here. That made us think there was something special about that class of compounds. We decided to pursue it further. I won't share most of the studies that we've done. We found that they bound to the same repertoire of receptors in the brain that haloperidol dose. They behave like an antipsychotic in a rodent model. This is a PCP induced model of hyper locomotion that is used for testing antipsychotics. You basically put rodents into a new arena. You watch them acclimate to an arena. They are injected with PCP, which, of course induces psychotic symptoms in humans. What you measure then is a hyper locomotion that occurs over the next hour. When you co-treated with an antipsychotic, you are able to blunt that hyperactivity and they just stay here at baseline. That is, in fact, what we see with these compounds that came out of this assay.

I think this has given us a feeling that we can really discover new compounds and a high throughput way. We can assess literally tens of thousands of compounds across a range of complex assays. Surprisingly, even things that you would not think would be easy to model in a zebrafish, like psychotic, antipsychotic activity. You are able through these rich phenotypic signatures identify activities as antipsychotics.

In closing, I wanted to highlight a couple of examples from other labs where other groups have taken a similar high throughput approach using behaviors to try and discover compounds that have neural activity. Here's an example from Robert Tanguay's lab at Oregon State where they used an assay called photo motor response assay. They took fertilized eggs, arrayed them in 96 well plates, allowed them to develop through the 24-hour post fertilization stage. They then exposed them to compounds, I'm sorry I guess actually beginning at six hours to 24 hours, at a range of different doses. They then exposed them to a behavioral assay that

consists of flashes of light at 30 seconds and 40 seconds. Then recording their response there. The normal response is the animals had some low level of basal activity. But after a flash of light, they had this robust photo motor response. They can become refractory and do not respond to the second stimulus. So, they run them through this high throughput behavioral assessment and then allow them to grow up for five days and run them through a battery of developmental assays where they look for morphological defects in these animals.

So they ran the thousand ToxCast chemicals at multiple doses through this paradigm. What they found is that about one third of the compounds showed very clear development behavioral defects at 24 hours. Most of those were characterized by decreased activity in the various aspects of the assay, but some of them caused hyperactivity as well. Then, they correlated those early behavioral deficits with the developmental deficits measured at five days and they found a high degree of predictability of developmental effects based on the early behavioral assessment.

In another example in Noyes et al. They looked at 44 flame retardants, putting them through a battery of behavioral and developmental assays at multiple doses. They were able to identify the different classes of flame returns could cluster into different groups. All 44 of them were found at some dose to have some kind of effect on either behavior or development. Interestingly, about 60% of the compounds had activity in all three of the assays. But they really needed all three different essays to capture 100% of the compounds.

Here is another example of a workflow developed by another group where they took fertilized eggs. They expose them, put them in 96 well plates, chemical treatments. Then they had a washout period of a few days. Then they exposed them to a behavioral assessment. After the behavioral assessment also eliminated ones that had died or failed to hatch or had other developmental abnormalities. The behavioral assessment that they were running was shown here where they were basically exposing them to periods of lightness and dark. What they found was that in the dark there was hyperactivity in the dark and reduced activity in the light. They were able to find compounds that had dose-dependent effects on one or more aspects of this behavioral assay.

A company that my lab has worked with and that I am a founder of called Teleos Therapeutics has also looked at a number of known neurotoxicants. They use a battery of 10 behavioral assays in an automated way to test the neurotoxicants at different doses. The results are depicted here. You can see the bars represent different things. In blue and red you have how similar the shape of the behavioral curve is to DMSO. As the behavior becomes less and less DMSO like, it gets more and more red. The size of the balls represents the amount of overall activity present in the wells. So, across a number of DMSO controls, you might see a little bit of variability in the shade of blue or the size of the ball. With the actual neurotoxicants, you can see a whole range of things, some that dramatically reduce the amount of motion. Others that increase the amount of motion and others that create a normal amount of motion, but the shape of behavioral curve is very different from that of DMSO treated animals. You can very quickly assess lots of compounds at lots of doses and graphically display the impact of those compounds on the function of the zebrafish brain.

In conclusion, I would like to sum up by saying that we think that systematic screening of zebrafish can identify compounds with diverse *in vivo* effects, including alterations in physiology. I showed an example of cardiac physiology. Importantly, also CNS function. The conservation of toxic effects of between zebrafish and humans is quite high. 95% in a number of cases that have been reported. Also, this phenocopy approach that I described is a way of

accurately predicting compound mechanisms of action. That certainly has been shown, I think quite well, with various drugs. Toxicants remain to be studied in greater detail. Finally, the system-level analyses can identify compounds with desirable profiles, if you're interested in drug discovery, or undesirable, if you are interested in screening toxicants, and to identify their mechanisms of action.

Let me just acknowledged the people in my lab who really contributed to this. I mentioned Dave and Andrew already had a lot of help from the other folks listen this slide. Barbara Caldarone at the Brigham and Women's Hospital who performed the mouse behavioral studies. And Andrea Velenich at Teleos who developed a lot of the analytical tools that I described today.

Thank you again for your attention. If there are any questions, if we have time, I would be happy to try and answer them.

Shafer: One quick question maybe?

Audience Question: [Inaudible]?

Peterson: So, the question was, how long do zebrafish live? What's their developmental timeline? There are actually two parts to that. There's development and then there's lifespan. Development is quite rapid. So after about five days after fertilization, these are free swimming larvae that are hunting food, performing all sorts of complex behaviors, sleeping, etc. It takes them another 2 to 3 months to grow to full sexual maturity. So, there is a [inaudible] growth phase. They can live as adults for 4 to 6 years, on the outside extreme. Development is rapid, but they have a fairly long lifespan.

Shafer: Thank you. We are scheduled for a break now. [Inaudible] Following that we will have the panel discussion.

Break

Shafer: [Inaudible] She does a lot of work in developing *in vitro* high-throughput assays for developmental neurotoxicity screening. She has worked with 3-D models [inaudible] to study molecular mechanisms using both human and rat models. She has also been participating in the development of adverse outcome pathways which our speaker following her will talk about. She received her doctor of medicine degree from the University of Düsseldorf. She did a three-year postdoc at the National Institute of Environmental Health Sciences in Research Triangle Park, North Carolina. I didn't know this until I read her profile here, but she and I have something in common. Our children were born at UNC Hospital. Anyway, Ellen, the floor is yours. Thank you.

In Vitro Approaches to Screening Compounds for Developmental Neurotoxicity Hazard
Ellen Fritsche, University of Düsseldorf, Düsseldorf, Germany

Tim, thank you very much for your nice introduction. I would like to thank the FDA and the SOT for inviting me here and giving you the opportunity to give a talk at this very interesting meeting and also, I am thankful for picking up [inaudible] important topic of developmental neurotoxicity

which is a very important aspect of children's health because it determines our society. We heard two beautiful talks already about developmental neurotoxicity in rodents and also in [inaudible] organisms. I will set the stage for a third aspect which will be the *in vitro* testing of the [inaudible]. We also heard some tips about their differences in rodents and humans. This is why my talk will focus on human models. My talk will be divided into parts. I would like to set the stage for giving some strategies and principles about the [inaudible] detecting. The second part, I will give examples what message do we really have? What can we do with them?

There is no conflict of interest here in my presentation. I would like to jump right away into the DNT testing. The principle of getting something complex like development of the brain into this is by taking this [inaudible] process apart, dissecting it into multiple processes that are necessary for building up a complex organism like the brain. This is a slide I really like to show. It is provided [inaudible]. These are the very basic processes of brain development like proliferation, migration, [inaudible] myelination, it and formation of functional network. One of these processes is heavily disturbed, we expect and [inaudible] outcome. This is why already going a little bit ahead [inaudible] pathways. All of these processes can be considered as key events that might be important in an [inaudible] framework. How we do this with these processes is this is the orchestration of the development of the brain. We take the process out, put it into a dish, and [inaudible] presence of compounds through the processes to act as they should without the compound in the dish? We're focusing on human methods, so do we had human methods where we can do this with for all of these different processes because the brain development is actually comprised of many different processes.

I would like to introduce, this is a publication that the European Food Safety Authority has initiated by sponsoring a literature review on methods that are actually there for testing development of neurotoxicity. This is across three different species here. Humans, rats, and mice. This actually tests different cell types, tumor and [inaudible] cell lines, primary cells, and stem and progenitor cell-based methods. Each little color here, you don't really need to read what this means, but it's one neurodevelopmental endpoint. So, in the middle here, you see the rat primary sales, this is a historical perspective. The most data has been produced so far on DNT testing *in vitro* has been performed on primary rat sales. Our target which will be the human here, you can see that actually the stem and progenitor cell-based methods are really just coming behind the rat primary methods, when you look at this in a matter of time, you can see that in the late 90s, the rat primary sales have really piqued. Now, really the stem progenitor cell-based methods [inaudible] really go up. This is taking publication from 1990 into consideration.

Building up on this publication which already proposed we might have an *in vitro* testing strategy here taking all the different major processes of brain development into consideration. The OECD has recently sponsored a report which takes publications until April, May 2016 into consideration. Focusing on which human methods are actually available to study DNT. The strategy of this report was first to identify been known AOP. [Inaudible] DNT have been published. That the compound-based mode of action. Will that [inaudible] compounds for which model, which signaling pathway do they act? And then in the first strategy, to look at basic biology, neurobiology by taking for example animal models, [inaudible] models into consideration. Which signaling pathways contribute to which neurodevelopmental process. Some genetic diseases in humans which [inaudible] development.

This was all collected in the report, and actually, these are just a few examples. [Inaudible] all of these contribute to [inaudible] brain development by actually interfering here or guiding certain neurodevelopmental processes. You see the different pathways guide different

processes. All of these processes are collected. All these pathways and then put into one big table, actually into [inaudible] graphical format because it is easier to read for us. So which processes do we really know are in the mind by signaling pathways we know. It is actually a lot of [inaudible] that we are he know. I am not going to go through each of these, but we need to differentiate [inaudible] from the stem cells, the neural progenitor cells, they need to propagate. We need to have [inaudible] neurons that have to migrate. We need to differentiate [inaudible], we need to make myelin.

So, this is going after neurogenesis [inaudible] all the way to [inaudible] network formation. [Inaudible] human-based methods where we can test of these. These are indicated by the little green hook. People have actually worked on methods where we can detect those. Then there are the little red crosses means there are cells systems there, but there are really no [inaudible] set up or any chemical has been tested on these endpoints. There is work to do. Here the little guys in brackets, there is indication that a [inaudible] will be ready soon but we are not quite there yet. Maybe with the rodent models, not for human folks. This is just to make it simple for you.

If we start putting each of these little colored boxes here into one assay and testing this, this will be so expensive, and it will take forever so where is the improvement to doing and *in vitro* by the way? The answer to this is you don't need a single assay for all of these because the, there essays out there to measure multiple endpoints. For example, here Assay4, this is the one I am working with. You can go all the way from the progenitor cells up to [inaudible] and neural differentiation where you can measure all this in one assay. You have maybe stem cell based and neurons where you can [inaudible] genesis all in one essay. So, we don't want to make it more complicated as it is. We are looking for assay which are able to do a lot of different things at a time and then combine these in a testing battery.

Talking about the assay. What do we want a assay to do? We can go over this. So, assays for DNT endpoint evaluations, they should have a high readiness level because we don't want to start developing another new assay. There is quite some [inaudible] on the market. So, it should be ready. The endpoint should be biologically plausible. It could be something that we know is functional and is necessary for brain development. And it should be human relevant. We have heard about this from Chip that cognitive functions in humans are often disturbed during development. So, we need assay, for example, AOP formation, we can protect the endpoints which are relevant for human function. So, then the cells obviously should be available. It should be something we have access to. The protocols should be standardized and ready to use so were endpoint specific controls for example have been tested and reproducible and model compounds are out there. We want assay that can do all of this.

Because we're talking about developmental process and we have embryonic phases and fetal phases during development, the process peaks during different phases of development. The *in vitro* testing battery is possible to really reflect this, so we should look at embryonic processes and fetal processes at early and later developmental processes. I talked about differentiation [inaudible] and on the other hand, as we look at the [inaudible]. These are very different processes that happen over time.

So, another thing is, we want the assay, if possible, they should be scientifically or biologically validated. [Inaudible] should be clear. This is an example here of the cell lines that has been heavily worked [inaudible]. These LUHMES cells, they differentiate into [inaudible] neurons, this is an example of [inaudible] inhibitor which is supposed to inhibit [inaudible]. It does actually in this assay.

Another example, this is some new data from my lab. The [inaudible], they can do a lot of different functional endpoints. When we look at [inaudible], micro-array analysis [inaudible] analysis, you can find all of these different processes. You find them back in the micro arrays. And then when you go for example that cell migration up here, when you go into the cell migration analysis, these are actually network analysis [inaudible] into gene function analysis. [Inaudible] central points and actually neurogenesis. As we look at our assay endpoints, [inaudible] inhibit migration without detecting viability, the same is true for [inaudible] inhibitor also interfering with neural migration in a assay, I will talk about this a little later. These are important physiological [inaudible] that are guided by pathways and the assay can really when you lack these pathways, the function is blocked. That is helpful if you have this information on the general biology on neurodevelopment.

We would like them to be upgradable, you don't really need to [inaudible] because you don't have [inaudible]. There is a low to medium throughput. If you have questions on the [inaudible], where you have thousands of unknown, you really need to [inaudible] for the zebrafish. There are different steps in a assay. A assay separation and then an evaluation of the actual outcome. So, the assay should be upgradable for both. Because [inaudible] work is really good 3-D systems, the neurospheres, I think 3-D is in some cases might get different results than in a 2-D system, this is already in 2012 can be upgraded the such a complex system to a higher throughput. When you have the right machine, this is a large particles order, you actually can. So, these are neurospheres. This is a comparison of the [inaudible] determination because we have a certain [inaudible] standard operation procedures, bimanual, and this is by the machine. So, it is doing this wonderfully, and you can plate them in 96 well plates. This is not cheap, but if you really wanted a 3-D system, you can upgrade it to a high throughput.

What about the evaluation of the data? This is actually, the US EPA has done a lot of work. These are some examples of their work for high content image analysis. This is a stem progenitor cell-based system which is an immortalized cell line that can be also purchased, this is a nice data that you can really investigate [inaudible] proliferation. And a high content analysis format, I talked about neuronal morphology. One is the [inaudible] cell, this is the stem cell-based [inaudible], so it’s purely neurons, it’s not a mixed culture. And also, the high content analysis can analyze neuronal morphology and the same is true with the LUHMES cells. The difference is really that here, [inaudible] identified measured in [inaudible] and center points, with the LUHMES cells, they are so dense that we can really only measure the total area that the [inaudible] cover. Different assays, different systems, different possibilities. Whatever you need, it is depending on your application. For synaptogenesis, there’s also, from the US EPA’s nice data that you can actually measure, I’m sorry, this has kind of moved in the what I cut out of the publications. But what you should see is [inaudible] that can be quantified in actually [inaudible]. This is highly sophisticated and will allow chemical screening for DNT for complex processes in a higher throughput.

Coming back to the 3-D system, this is actually a plated neurosphere where you have the neurosphere core in the middle, and cells have migrated and differentiated in the area which could be used as a high content system. We have actually written a software for analyzing this. This went online this week. So, on this website, you can download the software for free. I would be really happy if you would analyze [inaudible] with the software and give us feedback how it works. What it actually can do, it cannot only analyze [inaudible] but also [inaudible]. It cut out the [inaudible] core. And identifies all the nuclei, all the neurons [inaudible], it quantifies the neuron. It can also do the neuron morphology. That's why it can take whatever neuron pictures you guys have. [Inaudible] neurons and it has a learning algorithm in there which is

really easy to use. It can actually determine the migration area and determine the migration distance and also the neural positioning within the migration distance. So, it is actually geared to do high content analysis of a neurosphere. You can do high content with all sorts of systems and it can be adapted obviously.

So now this was the more general part. I would like to follow and go into the individual assay. The arrow that I showed you earlier, to set the stage, this is a beautiful picture of neural rosettes. [Inaudible] *in vitro* system of the [inaudible] formation where you can study compounds like retinoic acid which actually disturbs the formation of the [inaudible]. To help you where we are, the arrow will be shown up here. The red arrow will move across the big arrow so [inaudible] early or later processes of neurodevelopment. This is actually a [inaudible] satisfy two different lives. I know this is messed up [inaudible] neurotoxicity here.

What you can measure is one thing you can look at the [inaudible] in this assay. On the other hand, this is actually a sort of test taking only six days. Where you look at viability and the global gene expression. Gene expression, this is, up earlier one of my previous talks. I think it is a great [inaudible] for the future, but I don't think we're there yet. [Inaudible] genes are disturbed, and we can do risk assessment with it. This will be in the future, but we are not quite there yet. [Inaudible] gene function analysis. When is really the function of the process disturbed? Not the viability of the cells [inaudible] patent and see where our point of departure where there is no coming back to normal. I think for the functional assay, [inaudible] for the genetic modification, but I know that there is [inaudible] done by different lives on this.

These are assay that are ready to go. [Inaudible] getting a stem cell into the neural precursor. Then when we had the neural precursor for the brain to grow, we need [inaudible] ethanol. There are three different assay that I would like to present you were you can measure is such NPC proliferation. The one is the embryonic stem cell-based NPC proliferation assay. There haven't been so many compounds tested. This is rather a premature assay. It is a 2-D assay. Then we had the neural progenitor cells that you have seen now already. You can determine the growth and price over time. [Inaudible] two different measurements. We have tested not a huge amount of compounds but at least some compounds in this assay. There is an SOP behind this. There is the [inaudible] introduced earlier from the EPA. This is the assay with the highest readiness level. The 300 ToxCast compounds and a lot of other compounds including many model compounds have been tested with the cell system. It is also going in 2-D. You can do it a great high throughput with this test. Now, in this [inaudible] CX cell assay [inaudible] a second endpoint which was already done. [Inaudible] expression and activation in these cells and also a lot of test compounds have been tested on this endpoint whereas for the [inaudible] cells, this is more proof of concept study that you can measure activated [inaudible] and you can also detect [inaudible] independent [inaudible] that are [inaudible] positive which are these guys here but don't show [inaudible] activation in the assay. And also, of course, the embryonic stem cell-based [inaudible], you can measure in the stem-cell derived assay.

I am coming to the next endpoint which would be migration. Migration is really crucial during brain development. One example of a compound which interferes with neural migration is methylmercury. It is very well known for humans that it [inaudible] mode of action. We have actually different types of migration during the developmental process. One thing [inaudible] migration and for the neural crest cells, these do not only determine the brain development, but these are [inaudible] proliferate actually motor neurons, for example, but also [inaudible]. So, they come from the [inaudible] area. They also [inaudible] to other organs. The second one is the radial glia migration which is determined by the neuronal migration. So really three different types of migration.

When you look at this really closely, you can find more, but these are three basic classes. For the radial glia migration, this [inaudible] review [inaudible] that show that doing migration of all species which have the [inaudible] circuit of the brain [inaudible] we need to do other [inaudible] functions, you have different glia cells than in species that don't have [inaudible]. So, this is a depicted here. The basic radial glia cells. So, I really recommend this for reading. And then these guys here determine the migrating neurons which are [inaudible] who form the cortex.

What about assay which can actually mimic this migration? *In vitro*, one thing I would like to point out is the neural crest cell migration assay. A lot of compounds have been tested in this assay. It is actually based on the neural crest cells which developed from embryonic stem cells. These neural crest cells can be [inaudible] and you can do a scratch that this is actually coming from the development of a wound healing *in vitro* assay and then the cells migrate into this crest. So, this is very highly reproducible. This is a nice test which assesses neural crest cell migration.

The second type of migration, this is the [inaudible] that I talked about earlier. If you are lucky with a plating machine, you can also plate them in 96 well plates or in slides. It doesn't really matter. Then they migrate out, and this is mainly radial glia cell migration. And because these are human, and when you compare them to rodent [inaudible] cells [inaudible] neurospheres, we have mouse, rat, and our newest is rabbit, actually. So, in species where you don't have this folding of the cortex, you don't find these long radial glia cells in culture. You just don't see them. This is very specific for human migration.

This is how these guys look like. [Inaudible] red, [inaudible] after 24 hours, very early [inaudible] come out, they have this morphology here that really looks like radial glia in culture. You see up to 24 hours. This is a very quick assay. It looks a little bit different when you wait for five days. When you let them continue migrating and now [inaudible] in green, you get more like a [inaudible] maturation assay because you are close to, this would be the neurosphere core. While they migrate, they differentiate, and they maturate. So, you can look at a lot of different things by using a multi-cell type assay because [inaudible] neurons actually migrate [inaudible].

This is coming back to [inaudible] assay when you have the nuclei stained in blue here, the migration distance, this is after five days, this is why it is so big, can be assessed. We have tested all these compounds on the assay so far. For neural migration, don't get confused about the colors, but now the neurons are here in green actually. This is again the neurosphere migration area. When you take a picture. This is one picture of a [inaudible]. When you enlarge it, you see all the green neurons over time. I told you that the [inaudible] program [inaudible], how did the neurons migrate on top of the glia cells by dividing this into rings and looking at the number of neurons per nuclei in each ring. This would be the control situation, so they are divided equally. When you [inaudible] epidermal growth, that actually determines [inaudible] positioning in the [inaudible] cortex. You can see these are the ring numbers that the neurons [inaudible] with very low concentration of [inaudible]. They stop here in the rings, and they don't migrate. The radial glia migrate further, they proliferate, but the neurons release. So, you have a assay here will you can really look at how did the neurons migrate on top of the radial glia cells.

Coming back to the [inaudible] differentiation. So, there are different assay. Not too many available, but stem cell-based, stem cell-based assay where they look at the effect of ethanol

on [inaudible] differentiation. This is not an assay were chemical testing tested. Again, our neural progenitor cells, this is actually, from my lab, this is a different live used to this assay to look at the [inaudible], and they found [inaudible] differentiation. [Inaudible] induces maturation of [inaudible] *in vitro*.

So, another extremely important glia cell type and we know this especially from when you lose the oligodendrocytes when you are at old age from the [inaudible], there are not many *in vitro* assays that have looked at chemical effects on oligodendrocytes. But I would really like to point out this review here contains really interesting models that haven't been chemical tested through. These are models which can produce oligodendrocytes from embryonic stem cells and most importantly from IPS cells that could be possibly used for chemical testing in the future. Again, I would like to introduce what the neurosphere can do. [Inaudible] oligodendrocytes appear in green, and then red you see [inaudible] neurons. And oligodendrocytes occur in culture, as we usually measure after five days because then we have close to 10% oligodendrocytes in the culture. We can not only look at the formation but also at the maturation by looking at the expression of [inaudible] basic protein. This is after treatment with thyroid protein here which really compared to the controls accelerate maturation of oligodendrocytes because more myelin basic protein is produced. If you want to distinguish between effects that oligodendrocytes [inaudible] maturation of the cells, all you have to do is divide the myelin basic protein production by the number of oligodendrocytes and you get a maturation index, and this is actually right now under revision. There you see that T-3 stimulates maturation of oligodendrocyte. This is a flame retardant [inaudible]. I can tell you it reduces the number of oligodendrocytes in culture. When you look at the [inaudible] it doesn't interfere with the maturation at all. It is toxic to oligodendrocytes. You can distinguish between [inaudible] and compounds which inhibit the maturation.

Now coming to neuronal differentiation. You see quite a bit of work has been done from embryonic stem cells with methylmercury, ethanol, [inaudible], the bottom line of all this work is we really need a harmonization of [inaudible] here. This is very similar to what we heard from the [inaudible] study this morning. Everybody is using a different protocol for getting an embryonic stem cell to become a neuron or a mixed cell type culture. If we do not harmonize the protocol, we will never be able to compare any of these studies. I would like to point out, for example, there is one study that investigates methylmercury on narrow differentiation. You see the viability was down after more than 100 [inaudible]. But when you look at the [inaudible] positive cells which are the neurons, you get [inaudible], a different study has [inaudible] team expression decreases after [inaudible]. So how do you interpret this? I can tell you, these are lab centers using totally different protocols. So, I think protocol harmonization is a really important thing if we want to compare data across the different labs.

In addition to the embryonic stem cell-based test, the neurospheres also produced neurons, but you see, these look like the guys you saw on the slide [inaudible] migration. They are bipolar migrating neurons. This is a big difference to the embryonic stem cell-based [inaudible] because these guys you can get electrical activity. They are much more branched and mature than these guys are. We don't know why these guys are immature, but this is really an assay for the early differentiation into the younger ones and not looking further for maturation. And some compounds were tested also in this assay and endpoint specific control established.

Neurite outgrowth is one of the endpoints that has been probably studied the most extensively. There are different methods out there. One nice method is the hN2. This is also commercially available, and it's embryonic stem cell-based. It is a [inaudible] culture, and you can actually plate the cells in a way that the program can automatically test the neurite. You see the neurite

length increases nicely over time. You see the effects of compounds on the neurite length. This is also an embryonic stem cell-based but this is a mixed culture. You have [inaudible] and neurons in the culture. This is looking at the effects of methylmercury with also high content imaging. Low concentrations of methylmercury, and this has also moved, I'm really sorry. I'm trying to fix this on the online version, but very low concentrations of methylmercury actually reduce neurite length in the system. And it's mixed culture, so it has the advantage that you have to [inaudible] on the compound before it interferes with the neurons.

This is again part of [inaudible] because the neuronal identification algorithm that we're using is different from the one that the [inaudible] system uses because in our mixed density cultures [inaudible], we had a lot of false positives when we counted them by hand. One reason is the program of the high content imaging machine would say, this is a positive hit because the neurite is running along [inaudible]. And then it would say this is a neuron which has two branches. To make a long story short, our algorithm comes from here and tries to find a nucleus which is really in line with the neurite. And by this, we can get the false positive rate really, really much lower and also this is an example, this is the branching point. You see that the blue line, [inaudible] program actually has a different and actually the black is the [inaudible] accounted. They go one on top of the other while [inaudible] program has twice as much neurites. This is the reason because we have a lot of false positives in the culture. For the higher density [inaudible] there are solutions out now, and again, the LUHMES cell line [inaudible] basically. [Inaudible] or neurite length, but you can say okay, the overall amount of neurite in my culture becomes less. But also, this is a single cell type culture [inaudible] and there are no glia cells in there.

The LUHMES cells because they're dopaminergic, this is again [inaudible], they have developed a very nice protocol to differentiate them into dopaminergic neurons. So, this is the only assay that I know of so far where you can actually differentiate into certain cell types. This assay hasn't been used for testing either, but it would certainly, I think, be useful for this application.

Now coming to when you have neurons, you want to look at synaptogenesis and network and activity. This is very nice work from Tim Shafer. It is actually primary rat-cell based, so there is no equivalent there for humans right now, but this is the [inaudible] format, and there's a good protocol set up. [Inaudible] see how the neural networks work. [Inaudible] neurite outgrowth. When you see that the [inaudible] you can find nice concentration response curve when you grow the cells [inaudible] inhibits your development. I think the way to go will probably be having stem cells or preferably [inaudible] without any problems.

So, I would like to show this work hasn't been published yet. It was submitted a couple of weeks ago, where we actually made human iPS-based [inaudible] cells that also [inaudible] neurospheres and compared them to our primary neurospheres and showed that there are actually progenitors and when you differentiate them they make neurons and they also differentiate [inaudible], you find [inaudible] in these cultures and when you grow them on microelectrode arrays, you actually also get [inaudible]. The problem is, there are not that many, they are not that active as the [inaudible] that we grow in parallel. I think the standardization and reproducibility is an issue that really needs to be worked on. And also, other labs are doing this, harmonization of protocols again would really be preferred.

I showed you that there are assays out there. I would like to introduce something that was actually proposed for a meeting which will take place in Brussels next week. It's an OECD/EFSA, European Food Safety Authority, meeting where the idea is, do we have ways to

actually implement DNT alternative methods into regulatory application? And this is a proposal that might want to be discussed, a tiered-testing strategy for DNT could be envisioned by having pharmacokinetic modeling [inaudible] and then have a battery that I showed you [inaudible] as a tier 1 including alternative model organisms, and here the preferred organism is clearly the zebrafish. If you have a positive hit and then in the end, you need to go into a whole animal because what type of testing battery cannot do is [inaudible] or emotional functions, this is a clear imitation. If you want to go into rodent *in vivo*, maybe first check that your mode of action that you see in human systems is actually the same or similar to what you see in rodents by inducing rodent *in vitro* testing systems. This is crucial because if a mode of action on a similar basis cannot reproduce in rodents then in humans, you'll never see anything in the rodent *in vivo*. You spend a lot of time, money, and animals.

And why is this important? I would like to end with a couple of slides on species differences. This is actually early work from us why we actually started working with rodent neurospheres because we were looking for a hydrocarbon receptor [inaudible] modulator. [Inaudible] modulators are known in rodents, for example, to have an impact on neurodevelopmental outcome. We didn't find any effect of any [inaudible] modulation. So, then we went back and prepared [inedible] we actually did see [inaudible], for example, inhibited migration. When we prepared those [inaudible] animals, we did not see any effect at all. It was clear this was due to the hydrocarbon receptor in the spheres. And to have a long story short, when you look at the human, which is red here, [inaudible] there is absolutely nothing [inaudible] in the system. In the same year it was published that [inaudible] brains hardly expressed any [inaudible] development which is very different from rodents. So, molecules being present or absent during development can really determine toxicity of compounds.

This is an *in vitro* study again from the US EPA. This is a comparison of human versus mouse cells. This is proliferation of progenitor cells. I am not going into detail, but you see that [inaudible] that the responses of the [inaudible] compound is not the same across the species. And we've done this here for some endpoint of the neurosphere assay where we could show that human and rat neurospheres really respond differently to environmental toxicants.

I was very happy that Chip talked about the translating time. There is a web page that is called translatingtime.org that will compare certain aspects of development. Over time when we use this, our second trimester of gestation, and this is actually where our neurospheres are prepared from, correspond to, well it used to be postnatal day 5, but they've changed the algorithm, so right now, it's postnatal day 1. This is how we compare the timing of development in the *in vitro* cultures. When you look at the comparison of transcriptomes, the humans are the blue ones. The mouse are the green ones and the rat are the red ones. On the left, you see the proliferating neurospheres, and on the right you see [inaudible] differentiating spheres, especially for the [inaudible] that the human [inaudible] very far away from the two rodent species. And also, when it comes to differentiation, they are not really the same anymore. This is actually reflected in the [inaudible] blue is the human, red is the rat, and yellow is the mouse. So, the humans [inaudible] further away from the two rodents, but the two rodents are different from each other. We really have to take, this would be a take-home message, that the species different should not be neglected.

With this, I would like to sum up. From the current state of the science, the DNT *in vitro* testing can be based on neurodevelopmentally relevant processes that can serve as key events in an AOP-based framework. And then the signaling pathways to process function analysis so improve the confidence in the assay which is I think necessary for regulatory acceptance but also for prioritization. A large variety of neural stem/progenitor cell-based DNT *in vitro* assay is

available now but you cannot test individual assay. You really need the battery because if a compound inhibits the [inaudible] formation or the proliferation of neural progenitors, they are both equally important. You have to determine the mode of action. And similar to what we saw with the evaluation of the zebrafish, get a functional profile of your compound, understand how it works, have a library that you can come back to if you have an unknown, okay, the pattern is similar, my oligodendrocytes are really sensitive or [inaudible] or maybe the proliferation of the progenitors. By this, you can have fingerprints of your compounds across the functional assay. And then come to the genes later. So, the compound testing across a battery of *in vitro* tests covering timing and processes of brain development I think will be the next step forward.

With this, I would like to come to something that probably all of you know is the Creation of Adam. This is Michelangelo. So, Adam, he is there, but he is not awakened yet. So, in Adam, I see be DNT *in vitro* test battery. Here, it is God who awakens Adam. [Inaudible] testing compounds to awaken the battery to actually living.

The references I've used in this presentation, well, there are quite some. I would like to come to the acknowledgment because the OECD funded the Report on Integrated Testing Strategies for the identification and evaluation of chemical hazards associated with DNT. EFSA funded the Literature review on the *in vitro* and alternative Developmental Neurotoxicity testing methods.

With this, I would like to stop and thank you for your attention. I am sorry if I was a little bit over time.

Shafer: Thank you for a very nice coverage of a very large area. There is certainly a lot going on with *in vitro* assay, and you covered it nicely. Our final speaker is Dr. Anna Bal-Price. Anna received her PhD from the Institute of Pharmacology Polish Academy of Sciences at Cracow, Poland. She did postdoctoral research at Queen Mary College in London and [inaudible] University in France. She is currently a group leader working on developmental neurotoxicity testing strategies using non-animal approaches and human cell-based systems. She recently been very involved in the adverse outcome pathway development process. It really surprised me when we were putting this colloquia together that the AOP process has not been covered in previous colloquia that I could find. So, I think this is a very interesting concept. It is one that I think will be very useful in toxicology in general and in developmental neurotoxicology in particular. So, Anna will present us today the AOP concept and talk a little bit about how it can be applied and how data can be put into the context of AOP constructs. Anna, thank you.

Adverse Outcome Pathways for Developmental Neurotoxicity
Anna K. Price, ECVAM, Ispra, Italy

Thank you, Tim. [Inaudible]. I will go more into the detail and concentrate on these AOPs which are relevant to developmental neurotoxicity. So, no conflict of interest. My talk will cover introduction into AOP concept, some principles, how to develop AOPs. I will also present examples of AOPs that are already developed and relevant for DNT. I will finish with a few examples of how to use AOPs for regulatory purposes.

[Inaudible] ongoing transformation, moving away from [inaudible] towards more mechanistic-based regulatory decision-making. In this context, the AOP concept has been created, which entirely relies on the understanding of mechanisms of toxicity and hopefully will be used for risk assessment. So, it is a tool which help us in a very systematic manner organize existing

knowledge linking molecular initiating event with adverse outcome. In other words, mechanistic understanding what's going on with pathways of toxicity. And it very much underpins the paradigm shift in toxicology, moving away from this black box thinking towards building predictive toxicology.

One of the first AOP definitions has been published in 2010 [inaudible] saying that it's a framework which portrays existing knowledge and links molecular initiating events to adverse outcome through a cascade of key events at the cellular, tissue, organ levels, which induces an adverse outcome at the organism and then population level.

What is important in the concept of an AOP that identifies key events should be measurable and also, they should be linked to each other via causative or at least correlative manner. Usually, the definition of molecular initiating events, we describe here the knowledge which is behind the initial interaction [inaudible] gene and protein expression, changes in the signaling pathways, leading to the changes at the organ level causing deregulation of tissue [inaudible] function and development [inaudible] or interferes with the homeostasis process. And then these changes at the organ level cause the adverse outcome at the organism level [inaudible]. And obviously, changes at the organism level will cause changes in the [inaudible] population structure [inaudible].

One of the main characteristics of the AOP is that they are chemically agnostic. We do not try to describe what a single chemical does, but what any chemicals will do at the level of MIE. It is very much different from mode of action. We just focus on a chemical specific evaluation. What is important here is to describe what happens, the significance of perturbation of initial key events, which is MIE, leads to the [inaudible]. [Inaudible] will be applied in the predictive context. The, obviously, further information will be required on chemical-specific [inaudible]. And this features will dictate the strength and duration of perturbations at MIEs.

So, here is an example where all kinds of different chemicals, organophosphates [inaudible] would be relevant to the AOP. Not just one, but any type of the chemicals which comes from the same class of chemicals should give us the same interaction with a biological system.

So obviously, AOP is a pragmatic simplification of complex biology. You do not try to describe whatever we know about the pathways, but we identified key events. So, by convention, AOP is constructed by key events which are linked to each other in a [inaudible] manner, but in a [inaudible] way. That is really quite difficult to accept by scientists because we know that biology is quite complex [inaudible] responses, but anyway, for the regulatory purposes, we have to identify key events which are linked to each other [inaudible] manner, no branches. And AOP for a pure ligand is a functional unit of prediction and evaluation.

So AOP program of development and evaluation was launched by OECD in 2012. Then in 2013, the guidance how to develop this AOP and assess was published. It is very useful text because it covers all the different issues which are very important for developing AOP, what kind of information is required for describing key events [inaudible], how to evaluate weight of evidence. It also introduced some standardization.

Here I would like to mention the AOP [inaudible], which is a [inaudible] site for the submitting of AOP. We just develop according to OECD template. This AOP [inaudible] has been developed by US EPA and by European Commission. You will find all the necessary documents which will help you and guide you how to develop AOP. Recently there is also user handbook which

is a supplement to the guide. It is very useful because it gives practical examples to what kinds of information that is required for different parts of AOP development.

So, we start developing AOPs from the description of key events. So, here for description how does key events work? We should explain how [inaudible] changes in biological change are measured and whether they are really solidly supported by the current understanding of the knowledge that exist, so how well the key events which we proposed, is understood and documented in the scientific literature. This is the easiest part of the AOP, it's just not referring to the chemical exposure, just pure biological description.

Much more challenging is description of key event relationships, and here, we should describe the knowledge which is behind explaining whether [inaudible] is linked to key events [inaudible] in a causative manner. So, we have to first of all, describe [inaudible], so how well this key event is understood based on the biology knowledge, and then I said at the beginning, AOP is chemical agnostic, however here to prove the key event relationship, we have to use empirical support, but it does not mean that we create this AOP for the chemicals which we use for empirical support, it's an example of one or two reference chemicals which supports this key event relationship. Here, we should mention the experiment where it shown that by blocking the key event up, we also block key event down, and then we should provide some information to prove that the state of the key event up can predict the state of key event down. And it has to be strongly supported by biological plausibility, then we go more into the detail of empirical support.

One of the key information, if possible that we should comment, if possible, because it exist, showing what kind of threshold is triggering the key event up or down, that will lead to the key event down. What we see here is obviously the exposure to the [inaudible] chemical takes place, first of all, all kinds of events, mechanisms, will take place and adaptive changes, and having such a threshold allows us to define this concentration and this type of exposure, which is a tipping point where the cells cannot cope anymore because there cellular pathways [inaudible] strong enough, and it leads to the cell injury and toxicity. So, this type of information is very important. Obviously, we base AOP development on retrospective analysis of literature, so we take into consideration the experiment, they are not designed to support our AOP, so for instance, the context of the interior [inaudible] would very rarely have this type of information.

As I mentioned we have to use the chemicals which were started and supported in the key event relationship, and the other issue is the data should support the changes which are expected to happen in the certain [inaudible], so, key event up should happen first before key event down, but also key event up should occur at a lower level than key event down, and also, for given dose concentration, the incidence of key event up should be greater or at least equal to that of key event down.

The most important part of the AOP is overall evaluation of the AOP, and here, evaluation is based on the weight of evidence, and it is modified for your consideration, which are also taken as a base for quantitative evaluation in the mode of action, and biological plausibility is one of the first pieces of information showing that indeed the key events between up and down can be strongly supported by biological understanding. The experiments it's proved that blocking up or down, and the concordance of empirical observations, the dosage response, key events observed, temporality and incidents we can comment on the consistency, how well the pathway is preserved for different species, and whether the chemicals which belong to the same group, behave in a similar manner.

Since the AOP is a tool for regulatory purposes, we should see this in a transparent manner, so the only pick up the experiment and data which goes very well and we can support the identified relationship, here we also in a transparent manner identify adaptive knowledge, we detract from the understanding of the relationship, and also we should in the empirical support comment on these experiments which produces different data and do not support the relationship because obviously as you know if you write the paper, whatever he wants to support you always find the paper where it fits your needs so in this context we should be transparent to produce the documents which are objective.

Another issue after developing the single AOP, we had the next step which is networking of all of them together, which is much more representative to show complexity, another issue which is important from this network, that we can identify multiple molecular initiating events, which cause the same adverse outcome and vice versa. We can see the same MIE can cause multiple adverse events. So, from this strategy point of view which Ellen mentioned, we identify common key events which can serve as a base for *in vitro* assay development.

We can identify the events which are used by, which are present in different AOPs, and also, we can see the same MIE can trigger different [inaudible] down [inaudible] can be also triggered by different MIE, so it's just a next step in AOPs development to illustrate how complex the issue is. So AOP is are modular, key events and key event relationships are shared by multiple AOPs, so there is no need to rewrite the same key events that is important that if we describe key events it should be done in a general way, sometimes it is frustrating because it's not even allowing to mention it's possible to keep this general so a new AOP can be useful for other AOPs.

The network is facilitated by the knowledge base because when you do this, they are together. And obviously they are living documents, it would be updated with time as new data will come and you experiment so obviously it will never end.

Looking at the development we can be dividing them into three different categories, to take in AOP which are based on the events and relationships, and that is the case for the most part, they develop according to this international template, however, the AOP in this case is descriptive, there's a strong logical plausibility but you do not have enough data to quantitate quantitatively describe the relationship. And then the third category quantitative AOP, which has a solid data to support the key relationships and scientific confidence, it is much higher.

However, the good news is all stages of the AOP has potential utility, they will be used depending on the application, so it is very much the principal fit for purpose. So, for instance, the quantitative AOP should be good enough for screening of chemical categorization. However, if we already think about application of this AOP for the high level of confidence requires support by strong weighted evidence and such AOP has to be quantitative.

And the contents of the system there are very few pathways which are well documented, with outcomes so it is challenging, there's a general understanding of how a MIE can be linked to other outcomes, for example methylmercury exposure, it is linked to cognitive outcomes but one MIE triggers this, it is challenging. Also, physiological symptoms of the disease and vice versa, for instance autism is complex and defining the threshold for this key event up which triggers key events down and that's also very valuable.

Another issue is that DNT AOPs or neurotox AOPs [inaudible] really descriptive because we don't have enough data and then again taking into consideration of the complexity, we have to

move estimates from the singular AOP into the network. An example of an AOP, which we develop, this is binding of antagonist to receptors during brain development. It's one of the first AOPs already accepted, it's the process of reviewing and it is done at this stage.

We know this relies on the function of the receptor, activation, a direct link to the release, and for the construction of the AOP we review the literature after lead exposure, it is well proved that specifically of the research, it is higher at the areas of development and to illustrate what happens under the exposure, what is going on, under normal condition, that was quick. Okay. Under normal conditions, this is bound to the receptor [inaudible] this works with the receptor, leading to the release, at the end a triggers the reinforcement. In the presence of lead, which is bound to receptor, it leads to the decreased activation and then the lead activation and receptor takes place and decreases the release of the neurotransmitters and then this process leads to the impairment so that was quick what is going on.

The key events we have identified in this AOP, this causes reduced release and we have three processes which decrease this, and we have impairment of the memory, so I will skip this. I have an example of what kind of information can be used to support key relationships and information which gives you numbers and concentrations of the chemicals, which are linked to up or down.

There are other AOPs on the development which are relevant to DNT, we are finishing the AOP where at the molecular initiating event, we define inhibition of sodium iodide supporter [inaudible], which is one of the critical [inaudible] for thyroid hormone synthesis. At the end this AOP overlaps the previous ones because we have to development of the key events before [inaudible]. Similar AOP's are developed, one of them by FDA and the same others of the outcomes, it is reduced by another critical enzyme.

Here are the 10 that are listed I do not have time to talk more about them. I want to send you to the AOP site, you'll find all kinds information, so please go there and you'll see what kind of information is required for different parts of AOP development. One of the main aims is to encourage development so you can collaborate and help each other.

The few examples of how to use AOP regulatory context it depends on the maturity, we can create the assay which allows us to study key events of AOP, that would be to go over the screen assay to produce large data and this type of information can very much facilitate the predictive computation model, then the chemical categorization, this can be facilitated by AOP frameworks and one of the most important applications is for epidemiological studies, we can support the links and the language response for the studies and the risk assessment, for the high level and scientific confidence quantitative AOPs are required.

Here are references referring to DNT specifically, referring to AOP specific to DNT and general ones. I would like to acknowledge Dan, he is a part of our AOP training group. And also, Magdalini Sachana and Tomas Guilarte contributed very much to the AOP development. And I will stop here. Thank you for your attention.

Roundtable Discussion
Timothy Shafer, EPA, Moderator
Jason L. Aungst, FDA, Panelist
All speakers

Shafer: Thank you, Anna, very nice presentation. I'm sorry we had to go through so quickly. We're falling behind. I want to have time for discussion, so I want to invite the speakers to come back up for the panel discussion. I want to invite if there are people here in the room that have questions to go ahead and please ask some questions. We will have Jason Aungst from the FDA join us, he will represent the FDA and regulatory perspective. Any questions in the room? I can repeat the questions.

Audience Question: [Inaudible]

Shafer: Ellen and Anna, maybe you can comment on that, it has applications to the adverse outcome pathway, the question is when we look at these *in vitro* assays and molecular pathways, how much perturbation do we need to lead to an adverse, a permanent adverse outcome in the developing individual?

Price: The definition of this threshold is essential, then we can separate the defense mechanisms and changes from the moment when toxicity starts. How we came to this *in vitro*, I can imagine that we can expose the cells into the chemicals and then withdraw it for a few days and look at the same assay, if we do not have, see the changes after withdrawing of the chemical then we can say, the cells are still able to cope with it. I think *in vitro* in this case, is useful because we can trim design of the experiment to find the tipping point where we can identify concentration and time of exposure, where the cells were not cope anymore because the certain pathway you hit with such a strength, the cells are injured, the normal pathways [inaudible] because pathways of toxicity, so, it is challenging but I can see *in vitro* is one of the ways we can do it.

Fritsche: Maybe I can add on, another possibility but this only really works for compounds that you know are DNT compounds, is trying to find out what your internal exposure is, that you know causes a DNT effect *in vivo*. And then you can go back to the *in vitro in vivo* dose or concentration extrapolation and see what this concentration does in your *in vitro* system. And I think what is helpful if you had an idea not necessarily about toxicants but really how much of a singling pathway must be disturbed, for example, using pharmacological compounds, to produce a long-lasting effective *in vivo*? If we had information on this, and have a library, it can help us to go back when you come with the unknown compounds.

Shafer: I would add on one more thing, when you have data where you have both time and concentration information, there are ways of estimating the toxicological tipping point. So [inaudible] at the EPA has one publication and we're doing other work with him, so you can begin to mathematically define at what point does a system go from one where it can recover from an insult to one where there's a clear perturbation? So, but you need both concentration and temporal data to determine those tipping points.

Any other questions? You can use the microphone if you are here in the room.

Audience Question: The idea of responding, is this permanence, is this change, is that the cell itself has an internal mechanism to respond to this as opposed to, in a network of cells, what happens and the glia cells, other supporting cells that that whole milieu creates an environment where you can make a compensation, that is something we need a little more complexity on.

Price: Yes, I am concerned with the concentration and time of exposure, some can cope with it. But then, the cell cannot anymore deal with this type of toxicity and we can measure how to

distinguish this, this tipping point, to discriminate between the adaptive changes and toxicity and you mentioned glial cells which are extremely important as in supportive cell for [inaudible] that's why *in vitro* models we take the presence of these glia cells because for the same chemicals in two different models, [inaudible] as a pure model, and in the model where neurons are in the presence of [inaudible], the same concentration of chemical can induce completely different responses.

Fritsche: I can only support what Anna was saying, it depends on the application, if you want a real quick screen for, like, hundred thousands of compounds, to get an idea, I talked about the LUHMES cells and the measurement of this [inaudible] area, for example, that's a really basic screen but if you, but that will not be the correct answer so I think you have to go from throughput and complexity, this is the payoff you make. And balance of complexity of the system as a throughput and there was this nice diagram here with the zebrafish, it is giving the high throughput but has the complexity, but you have to make payoffs on either side and having more complex 3-D *in vitro* models that is somewhere in between, between *in vitro* and *in vivo* but it is human. You can never get more than 1000 data points a day, so you make the payoff on the number of compounds, so it is crucial you know what you want the system for, what's your intention? As soon as it gets into more complex and detailed signaling and pathways in human relevance, you need more complex models that contain all different cell types and that are preferably human. I think 3-D is superior to 2-D, but the throughput goes down, so.

Shafer: Any other questions in the audience?

Audience Question: I have a question for Dr. Vorhees regarding animal testing for DNT. There have been increasing advocates for group housing for animals. I was wondering from your experience, how much impact does group housing have on the DNT [inaudible], and what is your recommendation?

Vorhees: Animal welfare standards have changed dramatically over the last 20 years, now it is routine to do at least pair housing of rodents. And we know from studies that we've done housing conditions have a big effect, if they are done starting early in development, on learning and memory. We did an experiment once were we just added a stainless steel dome to each cage as the enrichment to meet [inaudible] enrichment recommendations, and that enhancement in the cage where they were housed 2 per cage, so they already had the enrichment of pair housing, and we added this stainless steel structure so they had a place to go in and out of, it had a big effect on their later learning and memory. So, that's one of the reasons why I'm skeptical of [inaudible] type observations like where you're counting urinations and defecation and things like that, some of those methods were developed when animals were put in single housed stainless steel wire bottom cages, nobody does that anymore, we took them out and handle them and put them down on a card that was stressful for them. That's not nearly as stressful now, when we do enrichment and pair housing and handle animals a lot, it is not the same measure that it used to be. And that's true of a lot of the observational things in the FOB, they are changing, and so had to keep these things standard, is very difficult but I think you can do it, now that we have some better guidelines on how to handle, house animals. I don't think anybody should be doing single housing, wire bottom cages anymore, I hope no one is, every cage should have enrichment, but if people start using different enrichments, inside the cages, that will create some issues. Yes, it is a complex problem. You still have your controls versus your treated, so you have within a single study, the controls, but across labs you will have some issues in that area.

Shafer: So, Chip, maybe you could comment really quickly on how that affects statistical analysis of data when you have group housing. I know there is a strong litter effect in developmental neurotoxic studies, so if you have group housing, how does it affect the statistical considerations you have when you analyze the data? And the numbers of animals you might need?

Vorhees: We always do a control for litter. We have the litter as a factor entered into the analysis in one way or the other, the most typical way we do that is put in as a random factor, since you do not control litter to litter variation and the offspring within the litter, if it is a postnatal study you can do a split litter design, if it is prenatal you're trapped, you have to do whole litters, exposed or not exposed, but within a split litter design, that is a powerful technique because can have all the treated dosage levels within the same letter. There are people who worry about that design, claiming, that the mother might detect which pup was treated and treat her differently, there's not really very, it's a hypothetical criticism mostly. The other thing, there is a new paper in nature that is interesting talking about the meta-genome and Microbiome and making sure you use littermates because that way you can make sure they have the same Microbiome. That is fine one litter will have the same Microbiome, right? What about across litters? Right now, we're setting up a study to try to make sure we keep the Microbiome the same across litters, by housing females together for a period of time before we breed them. So ,this is a new level of concern will have to start worrying about because it is clear the Microbiome has major phenotypic effects and in this paper they argue some of the lack of reproducibility biomedical, despite, in addition to not taking the litter into account using small sample size, etc., this is another level of control that we need to start paying attention to and we are starting to do that and I think most labs will have to do that.

Audience Question: A quick question, harmonization is very critical when you look at thousands of compounds, as scientists who are doing cutting edge research it is expensive, and difficult, obviously, for certain labs that may not have the funding to be able to afford to have the equipment that we have, so I am curious, in your respective fields, Dr. Fritsche and Dr. Peterson, is there a lot of sharing of data amongst different groups that do similar research as yours, and you have any suggestions or have you thought of ways to add harmonization? Because you have such, differences in lab financial capabilities, but is it also reassuring if you get data, tested the same compounds?

Fritsche: I think the last 10, 15 years, people have tried to get to a method to measure all these endpoints, there are different times obviously, in history, so we're trying to set up these methods that now we are at the point where saying when you make a status quo we have all these methods but now what are we doing with these protocols and what are we doing with all the different compounds? And the EPA has made an effort to reship the ToxCast compounds around so different people can actually, different labs can work with the same compounds with the same numbers that were attributed by the same people to try to start comparing so the new era is really harmonization of protocol and compounds, and trying to do some meta-analysis of the methods that have been developed over the last 10 years. I think it's now a new stage, so I think it's necessary but in the past, not that much, I have to say.

Price: I can also add that a few years ago we had organized together with US EPA the workshop where we set up the list of the criteria how to design the experiments for DNT evaluation, that is one step forward of standardization. So, for instance, design certain ways where you are playing with your experiments for acute or chronic treatment and how to define the hit, how to start from the standardization of the endpoints by defining the endpoint-specific controls. So, I think this kind of activities are going towards standardization of the approaches.

Peterson: I would comment for the zebrafish I don't think the harmonization has been good at all so far. I think, as Chip mentioned, even different lighting conditions, housing conditions can affect behavioral assays, so I'm skeptical that we will get really great harmonization ever and if you really want comparable in a high-throughput way, then I think it's best to have certain lab locations that do certain assays and do that in a routine way, that's the way to get the highest quality data, but as I showed, even if you have multiple labs doing different assays, once those become digitized into a fingerprint, they can be combined and give power to each other, even though they were done at different lots, so I don't think you need to necessarily harmonize it in that way.

Shafer: We have a question from someone on the webinar, and it's for Anna, Dr. Price. AOPs are linear sequence of events where each next event is dependent on the previous one, why there is a need to evaluate events that follow the MIE? Why not focus on MIEs only with QSAR [inaudible] targeted molecular assays?

Price: That is well spotted, you can use QSAR approach, for instance, for defining the chemicals which interact with MIE, you can use this MIE as the profiler. So, then, if you have a solid QSAR for identifying the chemicals that interact with MIE and in the AOP you already proved that once the MIE is triggered, this whole cascade of events takes place, and it leads to adverse outcome, one of the possible approaches would be to only identify chemicals by QSARs, those which trigger MIE and possibly in some cases that would be enough but you do not have to start the other key event if you can have strong biological plausibility and have a strong scientific confidence in the weight of evidence for key event relationships, then perhaps, you can also use this simplified approach.

Shafer: I agree. When you have a strong linkage between the molecular initiating event and the key events and the adverse outcome, then you can start to focus on those QSARs and the MIE as opposed to trying to test every level in between, but when you do not have that, it's important to establish those key event relationships.

Price: Exactly. And there's a lot of discussion, now whether it's from the regulatory point of view, if we used this key event as an anchor for [inaudible] development, whether it would be safer to go to for the key event which are close to adverse outcome, or whether this key event which are [inaudible] could also serve, that has to be shown by generating data.

Shafer: Any other questions in the room?

Audience Question: The zebrafish, you have all of the tubes, are they synchronized, or will it be in all different spaces when you're treating them?

Peterson: They are synchronized, so we take a lot of care to make sure they are all synchronized in their matings, and the animals all raised under identical conditions, so they should be synchronized.

Audience Member: Thank you.

Peterson: Thank you.

Audience Question: A lot of these methodologies worked well when you look at DNTs of single compounds. A lot of times, exposure tends to be concurrent with other compounds or

it’s a chronic exposure. Do you think these methodologies are robust enough that you can [inaudible]?

Fritsche: I’m not quite sure if I got your question right. Were you asking about the timing of exposure *in vitro*?

Audience Member: That and, the timing, if it's acute or chronic exposure, and if you're also, say, exposed to multiple pesticides at one time, or other chemicals, a mixture effect.

Fritsche: I don't really see the point why not if your *in vitro* system has a biological application domain to test different modes of action at once, that is crucial so if the mechanism is not there, then the second compound will not [inaudible] a different effect. If it works on the same pathway, of course, it's easy to see that but I think this is an important point because the application domain that you use the assay for has to be defined and I think we need a little bit more work on the assays in the future by looking at which pathways are there and especially which are absent so which affects can we not expect to see *in vitro*, that is crucial. And with the timing, it depends on the system, some systems can be kept in culture over long periods of time, for example the [inaudible] experiment, you need the neurons to be able to make network over time, and this is not quick, and in humans it takes longer than for the rats. So, our first activity we see after three weeks, they start but they are not really there at full blast, so for other assays, you don't really need to have the long-term exposure.

Peterson: I will just add on that I think, this combinations are a great thing to study and an important thing to study, when you start to think about combinations you can get to very large numbers of combinations quickly, even from a small number of starting individual compounds, and so having the throughput of these systems will be critical I think for being able to tackle that. I can say in the zebrafish we have looked at a number of drug-drug interactions that are documented in humans, things like [inaudible] that are mediated by liver, inhibition of liver metabolism of one drug by another drug, what we find is that those are replicated very well even in the early life stage zebrafish, so I think you will be able to find even complex interactions are mediated by different metabolic pathways where each drug alone has no impact but you put them together, in the case of [inaudible] erythromycin, neither of them alone has a strong impact on cardiotoxicity, but you put them together and they do have a profound effect and you can see that in the zebrafish.

Shafer: I would just add with respect to mixtures, one of the advantages of the high throughput approach is you can start to think about experiments that would be expensive and/or ethically difficult to justify because of the large number of animals that it would take if you have a complex mixture to really assess things like [inaudible] versus different departure from [inaudible] by doing those *in vitro*, if you have, as Ellen said, a system where the biological context is correct and you can start to look at a very complex mixture and look at each component of that and see where the biological contributions are coming from, so it's a powerful approach for mixtures.

Any other questions from the audience?

Audience Question: I had a question for Dr. Vorhees. The idea of the FOB, I agree with everything he said, it really does need to be redone, but it sounds like the FOB originally was used in the context of screening not as a definitive type of test. This whole idea of screening and using animals for screening, seems to be less important now with the introduction of all the other types of high throughput tests, *in vitro* systems. So, have we essentially done away with

screening in animals, or like you were saying, when they are doing the, when they were dropping the learning and memory tests out of the [inaudible] studies, was that the OECD, was that a, was their concept that that was a screening mode that the definitive test would be done including learning and memory in a developmental neurotox assessment?

Vorhees: I cannot speak for the OECD, but my sense was that it was actually designed as a combination to look at developmental neurotox and neuropathology and some functional behavioral effects, rolled into one study rather than doing the multiple individual studies. So, it's for that purpose, to investigate developmental toxicity, with various segments of animals going for different procedures, but I think they decided at some point that the test was sufficiently complex that they needed to streamline it some way, so they dropped the learning and memory. They wrote a review paper, looking at some studies in the literature, and showing that the learning and memory test was not effective, so they argued that made a case for dropping it out. I say, my view is that the problem is that the tests that were chosen, they looked at studies that use a lot of passive avoidance and T mazes, and I think that's not the way to look at the outcome. It's to fix it rather than drop it. So, yes, it's a screening, the extent of one generation is a screening procedure, it's an *in vivo* screening method designed to try to stop doing so many separate studies.

Audience Member: Is it still the option then to request additional data for more definitive dose response, dose, NOEL determinations and LOEL determinations? It seems like they do this one test [inaudible] acute study or something like that, and then they tried to use the NOL or the low effect level, no effect level, as the definitive means for regulating the chemical. Going to the next option of saying we need additional more definitive studies to be done. I don't know whether that's happening or whether, I don't know if EPA [inaudible] asks for additional data frequently.

Vorhees: I don't know how the extent of one generation study is being used in practice, that is not my territory. So, you'll have to talk to somebody who's in a regulatory agency to describe that. One thing I want to say, NTP has been working on a revised developmental neurotox battery and for those who are interested in seeing improved battery design, I would recommend you go to the NTP site, and look at their modified one generation and within it they have a, what is in effect a developmental neurotox battery, much like what I described that was needed. They have included pre-pulse inhibition, they have included the Morris water maze, they have done a lot more specification, so Dr. [inaudible], she is the one who's in charge of that, I think she's done a nice job and they're going to start using that on NTP compounds and we will begin to see data generated with this new revised, study design and see how the developmental neuro part turns out, and give us a better sense. And I don't know whether other agencies have been watching what NTP is doing. I don't know how much the federal agencies watch each other, but they are supposed to innovate and use those in the NTP program and hopefully that will have some influence as they collect data.

Audience Member: Interagency communication would be nice. Intra-agency would probably be better.

Shafer: The question here on the web for Dr. Fritsche. The tiers of DNT testing you showed appear to go from a human *in vitro* model to a zebrafish to rat *in vitro* and *in vivo*. What is the decision if there's a signal in the human model *in vitro* and in the zebrafish but not in the rat model? Does one assume no hazard? So, I think the question is, how do you start interpreting these data as you get them from these various different assays that have various different species represented and levels of organization represented?

Fritsche: I'm not a regulator, I'm an academic scientist, but I'll try to answer this question. If there is not a rat *in vitro* signal and no *in vivo*, I mean, we are human, so we don't really care for the rat, we care for the animal per se, but if it comes to hazard and risk, yes, the rat is a model like the cell model, so if one model does not show something but the other model does, I know in Europe they have hazard based risk assessments which has some rights but also has some flaws, I would go for the probable internal exposure, during pregnancy for the developing brain of the fetus, and if there is concern, if you want to go *in vivo* maybe look for different species but I don't think it makes any sense to go for rat *in vivo* study when the mode of action is not seen *in vitro* that you see in the human system *in vitro*. Then the whole thing to me, from logical point of view, not from the regulatory point of view, but to my logic it doesn't really make any sense, so yes, it's a good question but it was a proposal for discussion and I'm glad somebody took it up, but it needs to be discussed I don't have the answer.

Shafer: I think in part, it depends on the decision you have to make, and so if you're trying to evaluate, and I'll give you an example, when they had the gulf oil spill a few years ago, they were trying to evaluate which disbursements to use, and so you have a problem you have a set of chemicals that you have to choose from, and a number of the *in vitro* assays were run in the ToxCast program at EPA to try to find the disbursements that were the least biologically active. And those were the ones, that information was used in the decision-making process as to which disbursements to use. If you have a chemical that you're more concerned directly with a pregnant mother being exposed to in the diet, or something like that, then you will bring into the consideration the decision you make different types of information, as you said, exposure, does it get to the fetus, does it get to the developing brain, so I think part of the answer to that question is really, what decision needs to be made and what level of information and understanding of the pathways do you need to have to make that decision? I don't know if anybody else wants to comment on that.

Vorhees: Maybe I'm missing the question, but one of the things I think will happen is that you will have *in vitro* and zebrafish type signals where you get a signal and if you think it's a signal of concern, maybe take it into the rat DNT and maybe you get nothing, and you will have mismatches, then the question is [inaudible], and I think lead is a perfect example. Rats are not very sensitive to lead. You have to give a rat a lot more lead during development to get a clear signal out of a DNT type study than you do, we know humans are really sensitive to lead, and lead shows up as a signal in all the other assays but you'll have these mismatches and those will present some real problems when you get this whole system together. I don't know how to resolve them, it is a very careful case-by-case situation, I think, but one of the things you have to establish, it took a long time to sort lead, and you had to go from human to rat to even monkey, before the whole situation got clarified. Certainly we cannot do that with every potential neurotoxin, but when you get a mismatch, it will take a lot of careful thought and decide what is the most sensitive endpoint, and that's the one you will use to regulate, well, you use safety factors or however it is done, then you will take your most sensitive endpoint, regardless of which assay type it comes from, and then regulate from there using various safety factors. That's the only thing you can do. If the rat is insensitive, it is insensitive. I don't think you say, well, because the rat's insensitive, we don't have to worry about it.

Price: Like you mentioned, the current DNT testing is entirely based on animal testing, on the rodent, mainly, model, which is I'm referring here to test guidelines for, OECD test guidelines, and also, very much similar US EPA guidelines. So, even if we see something *in vitro* and we have proof that could be DNT compound, we still do not know what to do with the information because we do not have regulatory acceptance of *in vitro* methods, so that's why we have this

workshop next week where we take the regulators and the industry and stakeholders and the focus of this meeting is how to use *in vitro* data for DNT evaluation because so far were stuck, even if we have very solid information, we perhaps can use them for screening and prioritization but nothing else. So, I think we need some kind of decision-making process where we together with regulators would like to know what they need and what they can do with these *in vitro* studies and how to use this *in vitro* data. We have to talk to each other and agree on it before we can talk how we can use the *in vitro* data for regulatory purposes.

Shafer: So, we have another question from the web here. To the whole panel, with such a posse of assays and models, what is the minimal set of cell types and readouts that can be most informative for medium to high throughput testing? Who wants to take a crack at that one first?

Fritsche: I could throw a number into the room. I calculated roughly, depending on which assay you're using and how you combine them, I think the minimum would be five to six assays to really cover from the stem cell development to the neural progenitor up to the neural network. I think if you use them in a smart way, five or six should do, but this is just my personal estimation.

Price: I would say that it very much depends on the application, and I wouldn't be so brave in giving the number. It is very much depending of the aim of the assay and the testing strategy you are creating, whether it's to prove that this compound has a specific mechanism of toxicity, then you go to the very limited well-defined set of assays, but if you want to answer the question whether this compound has DNT potential, and you only know the structure of the chemical and nothing else, then I think it is much more complex.

Peterson: That's a tough one. I think an analogy for us is we are trying to use this battery of behavioral assays to figure out how to do drug discovery, and what we were able to do is start by measuring 14,000 different behavioral features, and that's what we started with. And then we were able to do the computational experiment where we drop out, so that was 10 assays and 14,000 behavioral measurements per compound, and then you can start dropping out assays and dropping out certain groups of behavioral features and see what happens to your predictive power and you can then actually get to a minimum number of data points and assays that you want to run, and you see dropping out certain assays hurts your predictive ability much more than others and you can start to do those kinds of, you know, just let the data guide you to what the minimum set is. So, I think ultimately, that's probably what you need to do in this field, you need to start with a big group of assays, and a good data set and start letting the computer model it and figure which assays are really giving you the most additional information and what the minimum set is that is truly predictive.

Aungst: I think one of the things we need to do, we have a lot of good data coming out, a lot of different assays available, so more of a reverse correlation would be necessary here where you take all this *in vivo* data that we already have, we're developing [inaudible], and then start tracking back to correlate what *in vitro* assays are necessary, rather than going from *in vitro* saying which ones worked, go backwards and correlate with the problem already.

Shafer: I would add my two cents, I think what you said, Randy, is very important, because we have a lot of assays, we have some data for many of them, I don't think in some cases we have enough data especially in terms of testing lots of chemicals, or compounds, not just the environment chemical space, and the entire chemical space, but all kinds of compounds, so I think we need to build, I don't know we can fully answer the question yet because I don't think

we have enough data to really allow the assays to guide us and tell us where, if we drop out an assay for proliferation, we really lose a lot of information here. So, I think, hopefully in the next five to 10 years, we really begin to develop that data set.

Peterson: I agree, but I also think you can pack in a systematic way. If you really say, here are 30 or 100 compounds that represent a nice coverage of the kinds of problems that we know are out there, and here's an equal number of compounds that do not, we know don't cause problems, now let's sit down and in a systematic way pick 10 or 20 assays and let's run all the compounds through all assays, you can then have a set that would allow you to come back and really in a quantitative, rigorous way, figure out what the minimum is.

Shafer: And there's also a couple of instances where those kinds of studies have been done yet. For a lot of the assays that Ellen talked about, we have maybe 10 or 15 chemicals at most that have been run in those assays and most of those were positive compounds with one or two or maybe five negative compounds, so the data needs to be developed.

Fritsche: The research for all these different assays, there are only two compounds which have been tested across a lot of these assays. One is methylmercury and the other is [inaudible] acid. So, these are the data-rich compounds where we have a lot of *in vivo* data for animals, for humans, and also for *in vitro* methods. I don't know, maybe people have a chance to look into the OECD meeting for next week, I will present a case study where I took all these assays that I presented here and researched what is the either statistical significance or [inaudible] value for methylmercury for inhibiting the different processes that I showed today, and what are the sensitivities? So that is interesting, I think something like this, this is catching up what you said, Randy, what you do with more compounds and then make a rank order, so which compounds [inaudible] mechanism, and which *in vitro* system is predicting what it should do? So, this is, I think, the way to go.

Shafer: I have a related question, [inaudible] published papers in 2011 and 2014, they described 11 total chemicals, 14 now, it changes, but between 11 and 15, the actual number is not that important in this respect. Those are the numbers of compounds that are known human developmental neurotoxicants, so what do we use as kind of a gold standard to validate all these assays including the animal assays against, when we have such a small data, a small set of gold standards, so I wonder if you want to share your thoughts? And then there is a question here for the panel, as well. So, maybe we could briefly discuss that and return to the other question, but it's kind of related to this. How do we prove these assays and what do we prove them against?

Fritsche: Coming back to the pathways, there are syndromes known, for example, mutations in [inaudible] receptors, that lead to an adverse outcome in humans. I think the way I would approach, besides these 12 for sure known DNT compounds, go for the pathways where we know there are human mutations that lead to human diseases, and we have [inaudible] of pathway modulators, inhibitors, and stimulators and include these in a battery to set the stage where we know this is the mode of action, we know there is an adverse outcome in humans, so if I disturb this pathway in my *in vitro* system, do I also see the adverse outcome? And this is important to generate a fingerprint library and from there go to unknowns.

Shafer: So, there's a question for the whole panel. Can the speakers share their thoughts on the utility of tissue chips and more complex systems than neurospheres or [inaudible] cultures? What will be the utility of the 3-D models and the more complex physiological systems in terms of testing chemicals and screening chemicals?

Fritsche: Yes, two things come to my mind. The one thing is the human on a chip project where, [inaudible] people are developing now [inaudible] where you have placenta models and brain tissue behind these placenta models. I think this will be very, very helpful. And the second technology, bioprinting, we haven't really heard much about this, but coming back to the variability of methods, I think the ability to bio print little, tiny bits of developing brain pieces and this in combination with the human on the chip technology, will be helpful because this allows us to produce models which are more standardized and more reproducible than the models we have right now. What has been developed over the last decade was necessary to get this technology, but I think in the end, this will be the future.

Price: I would like to add that the new technology is promising and to go for, but again, for regulatory purposes and screening purposes, we have to find the balance because maybe we do not need this extremely complex model for the question we ask, so it is not one answer but yes, that is the best model and go for it. I think even sometimes cell line is good enough. If I have a series of chemicals to find whether they activate an [inaudible] receptor, and I have a cell line and I know the that cell line expresses this receptor and responds in a way as *in vivo*, this cell line could be also good enough. So, we should have a balance between how complex the model are needed for certain purposes.

Shafer: I would just agree, I think the three-dimensional model and more complex models will be incredibly useful, especially if you want to think about patterning neural circuits and things like that, we cannot do in the two-dimensional model very easily. I think that will be very informative.

Any other questions on the web or from the audience? I have one that I'll throw out for discussion. Chip, you commented on how some of the batteries of tests that are done in the DNT guideline studies are not as informative as we would like them to be. I wonder if you and others might want to comment on as we develop these other alternative models, whether they be *in vitro* or *in vivo* more, do you think those could eventually guide what we might do in a *in vivo* study with rodents or something like that, and rather than bring this entire battery of tests that is currently consists of the DNT guidelines study, whether we might use that kind of information from the alternative test to select what we might want to test in terms of a behavior in a rodent model? And I wonder if you just would comment on that?

Vorhees: That's exactly what I would hope would happen. I think the example of the pathway for the [inaudible] receptor is a good example of how knowing if your compound affected that pathway, you would know some of the things to look at because you know that pathway is intimately involved in the hippocampus signaling for spatial navigation, so you want to look at spatial navigation in a test like that if you had a compound that sent that signal. So, I think that is an excellent way in the future to tailor DNT to the pathways that you can get out of whether it's a zebrafish or *in vitro* or an AOP it doesn't really matter, if you have some guidance, I think, you can get a good group of people together and make some more specific recommendations on how a particular DNT study might be done as your ultimate follow-up try for confirmation of the signals you're getting from these other systems. I think that's a great approach.

Price: Moving to *in vivo*, it would be extremely very well targeted experiment because you already produce so much information based on alternative approaches, and to confirm, perhaps you need only to go to *in vivo* for very well targeted design experiments to confirm, for the final confirmation, so also from the ethical point of view, this would really reduce further use of animals.

Aungst: I think you said about DNT, especially, like, OECD DNT testing, was a bit broad, a bit generic so far. But I think what you said gives a good opportunity, and we do this normally when we have a DNT study, we ask people to come in, talk to us about what they are planning on the DNT study, just not submit the standard study. Come in, let's look at all the information that is available, *in vitro*, *in vivo* data, other systemic toxicity data that may impact this, and then modify that study to fit the chemical we're looking at. So, it is being incorporated, and we're in a place right now where the area is so new and so broad and there is so much available, there is getting to be so much available information, we can modify that study.

Vorhees: I think that's the ideal way. I'm not, it seems to vary across agencies. My impression is, and you correct me if I'm mistaken, but my impression is at FDA there's a lot more interaction between the company making the product and the agency than maybe that goes on in other programs where I think a lot of data callings are issued and I don't know how much dialogue there is with the agency before the DNT is done. My impression is that the CROs take the standard DNT that they had been doing, then they take the new compound and they put it through, because if you look at what's available on the studies, they look remarkably undifferentiated in terms of the compound. They look like standardized approach. So, that may just be a difference between how the law and the agency allow for dialogue between [inaudible]. It sounds like at FDA, and I know this goes on in drugs, right? They will do a study whether it's a juvenile tox or different segment study, they talk and make a proposal of the design to the agency and the agency gives feedback, and the design may be modified. That's ideal. I don't know why it's a little different at EPA, but they may have to do with the regulations under which they operate, I don't know.

Shafer: I can briefly comment on that, although I'm not in the regulatory part of the EPA, maybe those initial 69 studies that you talk about in that paper that you looked at, maybe that was more the case, they just ran the routine battery. I know in more recent years, there have been discussions between the people in the regulatory part of EPA and the registrants in terms of are they going to do a DNT study, if so, do we need to do the whole thing, and sometimes they have even not done it and looked for other types of data, so there is discussion that goes on at EPA between the registrants and the agency as well.

So, we have a couple more minutes. I want to ask one question, I am surprised it did not come up at all. Maybe if everybody could very briefly address this. That has to do with metabolism, in the *in vitro* approaches, they typically have very limited metabolisms, the zebrafish have metabolic capabilities but how do those reflect the metabolism of drugs and chemicals in the human? So, I wonder if each of you could take a very brief moment to describe the issue of metabolism and how that affects the assays that you discussed. So, Ellen, why don't you go ahead, if you don't mind? Sure, or Anna.

Price: Metabolism is a bottleneck of any *in vitro* approaches, so we are very much aware that not many cell types are metabolically active under *in vitro* conditions. There are ways to deal with it, for instance, the addition of [inaudible] S9 is one of the way we can help with metabolism, however not in each case, because for instance neurons are toxic, they die in the presence of [inaudible] S9. So, what we're doing is we always tried to start from the chemicals we do not need to be metabolized to be toxic, or we directly use the metabolite. Obviously, we still struggle with it because we are aware that the toxicity in some cases might be induced by metabolites not by [inaudible] compounds. So, there are also *in silico* modeling which can help to predict the metabolite activity. Yes, that is true, that's the problem of *in vitro* studies, but I'm sure by combining *in silico* modeling and other types of systems which will support the

metabolism in the future, we will improve, and we will be able to test chemicals which toxicity is induced by metabolites. So far we have to be very careful to design proper the experiment taking into consideration this problem.

Fritsche: For me, as an *in vitro* researcher, my wish is that somebody is developing a good human liver and placenta [inaudible] *in silico* system, so that I can use now because I know *in vivo* the rat is not always a solution to the problem because drugs like ecstasy, for example, is producing different metabolites in humans than in rats and not only qualitatively but also from timing of metabolism, there are differences, so they are not always helpful but for me, a system that really could [inaudible] do good modeling of human metabolism, and really including the placenta would be crucial.

Peterson: Zebrafish have livers and kidneys and blood brain barriers and whatnot, so we have a number of examples where we have observed bioactivation, for example [inaudible] being active in zebrafish where it wouldn't be *in vitro*, I mentioned [inaudible] interaction, numerous of those where there is clearly an effective liver metabolism. I'm sure that there are differences, as there are in lots of systems, but I think it does benefit from having active metabolism.

Shafer: So, I guess one question I would have for you, and maybe just very briefly respond, is, how does the metabolism in the zebrafish differ from what we know in humans?

Peterson: There are groups that looked pretty closely at the specific isoforms of the cytochrome P450s, and there are many similarities, numerous differences as well, so probably if you're trying to drill down to the exact cytochrome P450 that's metabolizing a drug, you may not find the exact [inaudible] there, so I think more work needs to be done to figure out what all of the differences are, but at a functional level we see the same kinds of metabolites being produced.

Vorhees: I would just say, the problem of metabolism, as has already been alluded to by Dr. Fritsche, is a problem in rodents just as must as it is in any of these other systems and that sometimes the rat does not metabolize the same way as people or the mouse or, you know, we are always facing that problem of the metabolites [inaudible] internal dose versus humans.

Shafer: Thank you. I think that is all our time. There are a few announcements. I want to announce a Blackberry, for the people in the room, was left on the table during the coffee break. If that is yours, you can pick it up at the registration table. And then, Jeff, did you have some closing comments? I want to thank all the speakers today. They did an outstanding job. I appreciate all the time and energy and effort that you took in developing your slides and I appreciate you being here to present them, so thank you to all of you.

Jeffrey J. Yourick: I want to thank Anna Price, Ellen Fritsche, Randy Peterson, and Chip Vorhees for readily agreeing to speak today, it made our job much easier. Thank you again for doing that. I want to thank Jason Aungst, from the Office of Food Additive Safety, for being on the roundtable discussion panel today. And Tim Shafer, we owe him a large amount of gratitude for putting this together essentially and really spearheading this program today, so we all appreciate that. Also, I want to thank Betty in the back, this would have been impossible without her efforts, and the rest of the SOT staff and the FDA staff college for their assistance.

I also want to thank Allen Rudman for being a [inaudible] behind this whole program and Ivan Rusyn for chairing. So, with that, I just want to make everybody aware of the upcoming colloquia that we have scheduled. You can see on the screen on December 1, we have the

Application of *In Vitro/In Vivo* Extrapolation in Safety Assessment. That should be a good program. Look for those announcements to start coming out fairly soon.

Also, in the spring, and we're working on the timing, but we might piggyback this on with the SOT in Baltimore, it's still under discussion, but we will present a colloquium going over the report that's just been worked through and approved using 21st century science to improve risk-related evaluations. And sometime in the spring of 2017, we will look at clarifying adversity in food safety. So, look for that later in the spring.

Also, please go to the website listed on the bottom of the slide to look at this recording as well as recordings and slides from all the previous colloquia. Use that as a resource. They're there. It is free. Tell your coworkers. Tell everybody that you would like to. That's a nice resource.

Please, thank you to the participants. Thank you to the participants in the room. Thank you to the participants that attended on the webcast, I know it's probably a worldwide audience, and we appreciate and hope to have that continued involvement from everybody in the future.

For the FDA staff that are here, we will be having an informal luncheon with the speakers directly after this. I'm not sure of the room because that's been changed. 1B 042. So, please join us. It's bring your own lunch. Thank you again to everybody.