Risk Analysis of Endocrine Disrupting Chemicals Using Higher Animals

Yasuhiro Yoshikawa

Graduate School of Agricultural and Life Sciences, University of Tokyo

Thank you for the introduction. Our team is a part of the Kuroda group. We were allotted the themes of how to conduct risk analysis of endocrine disrupters on neural development using higher animals and how to extrapolate those results to human beings. Our research over the past year and a half or so has primarily involved rats. Experiments with monkeys are to begin in full earnest next year, so today I would mainly like to give a description of our basic strategy, including the data we have obtained up to now.

Consisting of Drs. Yoshikawa, Negishi, Yasumoto and Suzaki of the University of Tokyo conducting fundamental research involving chimpanzees, monkeys and rats, Dr. Kawasaki of Hoshi University conducting research on cautionary behavior of rats and learning analysis of monkeys, Drs. Koyama and Shimomura of Japan Women's University conducting behavioral analysis of monkeys, and Dr. Kuroda who is in charge of the group, our group is a rather hetero group. The object of the research is to extrapolate analysis results from rats at the very forefront of the slide to monkeys, chimpanzees and human beings – I do not represent the human race.

As for today's topics, I would like to introduce the following three points:

1) How to extrapolate ADME (Aspiration, Distribution, Metabolism, Ejection) of endocrine disrupters beyond specific difference.
2) How to extrapolate neuro-behavior, the main theme.
3) The fact that cooperative research including laboratory studies and field work involving Japanese monkeys is needed.

This slide shows bisphenol A (BPA) as a target. The slide shows transition of blood concentration in various animals after injection of BPA. The top gives an example of oral administration of 10 mg/kg. When 10 mg is administered orally, BPA concentration can barely be detected using the ELISA kit (kit for measuring nonbonding BPA using monoclonal antibodies). This is indicated by the blue point in the figure. Four rats were used for each measurement point.

Compared with this, BPA migrates in the blood of monkeys with extreme efficiency. The figure for monkeys shows the average for groups of three monkeys each. Each group of chimpanzees consists of two individuals. Data is given for two individuals each. The figures for chimpanzees fall right between those of rats and monkeys. The results are the same for subcutaneous administration of 10 mg as well. The substance migrates with greatest efficiency in the blood of monkeys (*macaques monkeys*). Chimpanzees tended to be somewhere in the middle, with rats being the lowest. The tendency was the same for rats and monkeys at the high dose of 100 mg/kg (chimpanzees were omitted because the dose was high).

What this means is, in order to extrapolate generally beyond specific difference, risk analysis is enhanced by safety index, and it may not apply to a simple phylogenesis in all cases. I think that this must be taken into consideration. I think this phenomenon naturally differs according to the type of chemical substance and I think there are specific differences such as degree of migration into the blood, actual

metabolic enzymes, CYP 450 and difference in metabolism due to conjugation enzymes etc.. I must point out however that when initially extrapolating dose, the point cannot be raised by simply enhancing phylogenesis by a priori.

We have observed reaction of single dose with comparisons so far, but there are various limits applied to experiment design when using an actual human being as a model. You naturally want the experiment to simulate conditions of human exposure as accurately as possible. In the case of routine exposure, for substances that are so quickly metabolized such as BPA, the substance would be administered on a daily basis consecutively. In the case of rats, the gestation period lasts 21 days and the nursing period also lasts 21 days, so oral administration on a daily basis is not so difficult.

In reality it is almost impossible to do the same thing with monkeys. For monkeys (*macaques monkeys*), the average gestation period is 155 days and the nursing period is about a half year. The results of the test becomes distorted due to stress if we try to forcibly administer a certain oral dose by probe to the mother every day. In order not to put so much stress on the mother, we can get by with subcutaneous administration by osmotic pressure pump once a month instead. In this case a capsule is buried under the skin.

If you use the extrapolation mentioned a little while ago, the index for rats and monkeys is about 10 times and migration in the blood is 10 times for oral and subcutaneous administration regardless of type, multiplying by index of about 100, the experiment with monkeys would begin with a dose of about 1/100. In the case of chimpanzees, besides extrapolating by the single dose mentioned a little while ago using healthy animals, the experiment is impossible in terms of animal ethics. Experiments involving long-term administration are not possible at the present.

Concerning persistent substances such as TCDD on the other hand, it is possible to administer oral doses once a week or once every two weeks for rats and even monkeys, so there is no difference in methodology and it is very easy to extrapolate the results to human beings. In experiments with chimpanzees, there has never been a case when even single administration of TCDD was allowed by the animal ethics committee, so extrapolation from this level is impossible. In any case, the first step toward measurement of neuro-behavior requires basic analysis of each individual substance to find out at what value should dose and method of administration be extrapolated.

Secondly, as for the difficulty of analyzing the neural system, particularly neuro-behavior, compared with endocrine system tissue, there is an extremely large difference in terms of phylogenesis for the neural system. The slide shows the brains of various animals, but it is not easy to extrapolate by simply enhancing phylogenesis. For example it is very difficult to extrapolate in the same manner for animals with highly evolved neocortex such as mammals, animals with striatal exposure such as birds and fish, corpus allatus (*corpora allata*) for which terminal of the neural system has become large and insects and invertebrates that stop evolving at the mushroom body level. If you line up Primates that are the closest to human beings, especially old world monkeys chimpanzees and human beings, as you can see just from looking at the slide, the size – as is often said – there is an extreme amount of difference concerning development of the prefrontal area. There is a problem that something such as this has to be extrapolated.

Concerning assessment of neural system function, you want to basically carry out assessment in levels of several stages. The slide shows the most basic circadian rhythm. I would now like to talk about

rats. Rodents are nocturnal, so their behavior pattern clearly peaks during the night. Most of their activities decrease during the day, and they are often either sleeping or dormant. This is the mouse system. This is a mutant mouse we discovered, which exhibits abnormal behavior and emotion. We found that there was an abnormality in the genes related to kinesin. As the slide shows, the circadian rhythm of this mouse is out of whack. If the rhythm is out of whack day and night up to here, the result is clear even if measured, but the difference is not that pronounced for substances such as BPA.

This is the behavior for a single day for the offspring of a rate orally administered 4 mg, 40 mg and 400 mg doses of BPA during pregnancy. If we look at the activity of rats during their waking hours, in other words, during the night, we find that the dormant time, the time where their activities have ceased, at 40 mg/kg for the group of 4-week-old females only is significantly high. There is sex difference in the effect of BPA here. We are not seeing a simple capacity reaction, but rather the complicated result of losing this tendency due to development.

The level of the second stage is an unfamiliar environment; there is the problem of how to induce behavior, especially when placed in an unfamiliar environment such as an open field. In the past video taped behavior was often recorded with paper and pencil by people. “Cineclusters” (brand name of a type of motion picture) were developed when we tried to automate this entire process by computer. As a matter of fact, the very first one went on sale last month. The price is $20,000 for a complete set. With this system, behavior video taped in an open field is converted to digital format, separated into elements on the screen and calculated. The software is equipped with an algorithm that decides how to separate different types of behavior. Digitally converted images are automatically extracted from various screens and center, length, breadth and topographical gravity position data is analyzed. The results are extremely simple. The system displays the results as a color bar code like you see here in the slide.

These are the results of time series analysis of open field behavior of a rat analyzed by Cinecluster. At the bottom are the behavior classifications analyzed by a professional rat behavior analyst: changing location, rising, stretching, lying dormant and sniffing. Behavior is separated into actions using the whiskers only, actions requiring head movement, face washing and grooming. The rate at which the results of Cinecluster and professional analysts agree is about 70%. Concerning the offspring of rats exposed to BPA during pregnancy, an open field test was conducted on groups of male and females 4, 8 and 12 weeks old. The results were analyzed by Cinecluster. As you see here, grooming appears to have increased significantly for 8-week-old males administered a low dose of 4 mg/kg/day.

The next higher level is learning and memory. Existing stuff was used for the water maze and shuttle box test (active/passive evasion test). The slide shows the results of the active evasion test. Understanding tended to be significantly low in session No. 1 of 8-week-old males of the BPA administration group. This may be related to the results of the open field test (significantly increased grooming). It could be that the capacity to react to external stimulus has deteriorated rather than capacity to learn being poor.

Level No. 4 is the drug administration test. Up to now, unfamiliar environment or learning capacity has been studied after allowing it to roam freely for a period of 24 hours. Even this is however a covert behavioral abnormality. The threshold values for capacity to react to stimulus are unusual, and so not become apparent under ordinary circumstances. For example, there is the problem of how to make fine function impairment obvious. The test that takes this into consideration is called the “drug induction test.”

With this test, the specimen is administered a stimulant such as caffeine, amphetamine or L-dopa to see to what degree behavior is altered and whether there is discrepancy in the threshold setting point.

This is a comparison of open field behavior. Behavior is completely unlike that of a normal litter when administered caffeine, 10 mg in this case. If the elements of this are analyzed, hyperkinetic and excited movements such as moving about or standing up become more frequent. On the other hand, behavior such as sniffing and grooming becomes less frequent.

Now let me switch the subject to monkeys. With this project, we conducted the same experiments we conducted with rats, this time using monkeys, and compared the results. Monkeys' behavior is much more complex than that of rats. This shows six patterns. Classifications of behavioral scientists that specialize in monkey divide behavior in 14 patterns. There are several basic actions including turning around, standing up, and sitting down. Just as with rats, the motion picture processing extracts the monkey from the background and determines the movement finite difference, total area, major/minor axis length and center of gravity. Concerning less frequent movements, movement of parts is *merkmal* and an algorithm is created.

This is the result of analyzing the four most basic movements using the algorithm. In this case there are four elements, and at 92%, the rate of agreement between computer and observer is extremely high. Because a monkey's actions are complex, with as many as 14 patterns, we have still not been able to come up with an accurate algorithm.

After a pregnant monkey administered an endocrine disrupter has given birth, we begin to analyze the interaction with the mother. This is done by videotaping six months following birth. The behavior of the mother toward her offspring includes stereotype behavior such as holding, grooming, and lip smacking. Oppositely, we decide other items of behavior of the offspring toward the mother such as touching the abdomen, touching the back, arm/leg contact, getting upside down and so on. We also decide other items of behavior of the offspring apart from drinking milk itself, climbing on the back, sleeping and making noise. How do these items manifest during the developmental process up to weaning and in what way do they shift? By dosing so we find out how these items differ by exposure to substances such as BPA and TCDD.

In order to determine the social characteristics of monkeys, pairs of baby monkeys are raised together after passing the age of six months. This experiment is the equivalent of the open field test for rats where rats are made to socialize in an unfamiliar environment. If anything, the behavior group that positively acts toward other individuals and the "self-directed" behavior group that directs behavior toward itself are given as items. This experiment is conducted over the period of about six months in for ages of about six months to a year in various combinations such as BPA-BPA group, BPA-control group and control-control group individuals.

The third level is the intelligence test. The slide shows the finger maze test we developed because we needed an intelligence test that can be conducted by anybody anywhere and that can pique the interest of any type of monkey. A four-stage finger maze is currently used for the test. With this test we place apple slices in a box with slits. Monkeys like apples so they try really hard to get them as soon as they discover them. The monkeys can stick their fingers in the slits and maneuver the apple slices. If they move the slice in the wrong direction, they can't get it because it is an error box. They can get the apple only if

they move it in the correct direction. If they pass the first phase of the test, they advance to the second phase. In the second phase, the direction is unfairly reversed. If the monkey tries to use the same strategy that allowed him to pass phase one, he suddenly goes to the error box. He must discard the strategy of the first phase and go in the opposite direction. The third phase calls for switching back to the original direction, and reversing again for the fourth phase.

This slide shows the test in progress. When performed on human babies, the babies under three years old are generally unable to pass the test. Monkeys mature more quickly than human beings, and can pass the test at age one to one and a half years. For the first phase, for example, the test is conducted twice a day in two sessions for a total of 15 trials. If the monkey passes two sessions of 14 trials consecutively, he advances to the second phase and all the way up to the fourth phase. In the joint research conducted by Dr. Yasuda's group, the test was conducted on baby rhesus macaques born to mothers exposed to TCDD. The data is now being arranged. Why the group exposed to the highest dose of 300 ng had the highest score and passed to the fourth phase the quickest in contrast to expectations still has us scratching our heads. We analyzed social characteristics of the encounter test and scores of the finger maze test and analyzed correlation with items such as introversion by the eye contact test. We now plan to assess influence of TCDD.

The fourth level, as we explained when talking about rats, is the drug induction test. This is an experiment conducted together with Dr. Inoue's group whereby methamphetamine was administered to rhesus macaques. We are now considering whether to analyze behavior of specimens administered several drugs including L-dopa, caffeine and cocaine by assigning a task or analyze free behavior using automatic analyzing equipment in the same way as was used for rats. Thus we can lift and extrapolate behavior level assessment of rats, monkeys and human beings. Circadian rhythm serves as basic element of behavior analysis. In the case of human beings, the pediatrician would ask the child what time he got up and went to bed yesterday throughout the year.

The unfamiliar open field environment test, in the case of human beings, would be to bring an infant to the park for the first time. In the case of human beings, we have accumulated an extremely large amount of data on mother-child behavior. The encounter test is the equivalent of bringing children to a kindergarten where they interact with one another. For learning and memory of rats, there are shuttle avoidance and water maze tests. For monkeys, there is the finger maze test. Monkeys are capable of passing this test at one and a half years of age. Human beings don't mature as fast as monkeys, and are capable of passing this test at three years of age. Then there is the drug induction test. This test can be conducted using rats or monkeys as specimens, but may not be conducted using human subjects. I think it would be difficult to experimentally extrapolate. Thus comparing behavior assessment obtained at each level with human beings, the objective is to clarify the effect of endocrine disrupters on neurological development. Of course, behavior abnormalities assessed in this way are a result of abnormal development of what part of the brain is connected with research on the histological and cytological level, and the ultimate objective is to clarify this mechanism.

Now let us turn our attention to the third topic. Up to now we have been talking about strategy of experiments conducted at laboratories. This is the main function of the wild monkey. As you know, Japan is an advanced country and is home to few varieties of monkeys. It is the home of the northern-most habitat of Japanese monkey. Here and there we have wild monkey parks for tourists where the monkeys

are fed. Experiments concerning social behavior of these monkeys were conducted, and we found a culture of other than human such as potato-washing monkeys.

At the beginning of the 1970s in Japan, deformations in monkeys living in wild monkey parks, especially deformations of the extremities, were all together observed becoming a major problem. Because this did not occur in monkeys living in the wild, research was focused on groups of monkeys living in wild monkey parks that were provided with food during the period of 1970 to 1975. The Ministry of Education, Government of Japan therefore quickly assembled a research team and analyzed the problem from various aspects. In the beginning, there was a certain degree of family gene accumulation, but as a result of genetic analysis including the mating experiment, single dominance or recessive genetic material was negated. Various items such as nutrition, poison and viral infection that causes special deformities such as rubella were studied, and these were also negated. Finally, the first candidate raised the possibility of pesticides in the food. The imported feed used at the time was changed after which deformities did not appear. Research was therefore ceased.

In addition to wild monkey parks, wild Japanese monkeys are distributed up to the Shimokita peninsula, so their habitat may be near places where humans reside and therefore have many opportunities to come in contact with pesticides and environmental chemicals. If we can conduct joint research including laboratory and field work and the results are tied in with research on the cell and gene level like the previous person was talking about, the result would probably be more significant.

There is still a lot of room for debate concerning whether the endocrine and sophisticated neural systems of higher animals are strongly affected by environmental chemicals. Endocrine disruption began with steroid imitators, and as the previous speaker mentioned, substances with a similar chemical structure to thyroid hormones could possibly affect neurological development. As I said today, however, it is not so easy to assess risk. If you try to systematically raise phylogeny without analyzing individuals in particular and you don’t organize your strategy, you will end up with a singular result rather than an overall view of the situation. It would also be impossible to analyze risk if you don’t approach the task in a comprehensive manner involving a system of cooperative research including field work as well as laboratory work.

That’s all the slides I have. Thank you for your attention.

Q&A

Koibuchi: Thank you very much. The paper is open for discussion. Any questions? Yes please.

Q: You stated a little while ago that the monkeys performed better at 300 ng in the TCDD. The mother monkeys were exposed, weren't they?

Yoshikawa: Yes, the mothers were exposed.

Q: I have heard talk that it is possible that the monkeys performed better due to increased social contact with the exposed mother.

Learning generally, with radial rays of rats for example, reference from the memory ceases to exist if the hippocampus is destroyed and the specimens no longer become confused by memory referral, and in some cases, there have been reports of improved performance as a result. But although performance is improved, does this necessarily mean that dioxin is not bad? For the same reason, this is being conducted at level 4, but when difficulty is gradually increased to levels 5, 6 or 7, I think a different effect may be observed. What do you think?

Yoshikawa: It is quite possible that the results may be influenced by the initial problem, maternal care. Previous experiments of this type had already been assessed. The monkeys were already over one year old when we got them and took the measurements, so we do not have that data. When we pursue that with BPA, we are thinking of getting it (the data) from there.
As for the part about rats, I think this may be extremely helpful. What I myself actually developed was rather this system for measuring an Alzheimer model in aged monkeys. If the specimen decides on a certain strategy and adheres to it, the first level will be exceeded. The fact that second level however requires an awful lot of time but the fourth level can be passed quite quickly does not necessarily mean that it is illogical, but rather the specimen cannot find his way through the finger maze unless he has a flexible strategy. When the level is subsequently raised to levels 5 and 6, it takes such a long time to get the reward, many of the monkeys give up before they get there. We put the system together gradually from level 1. Reproducibility is generally highest up to about level 4, so we are currently using a 4-level finger maze.

Koibuchi: Dr. Rice has a comment.

Rice: I just wanted to comment on your TCDD findings and the finger maze. As I am sure you know, Sue Chance, when she did her dioxin study she also found facilitated performance in some kinds of discrimination tasks, and impairment in others.
Her group interpreted that at the time as being indicative of the fact that control monkeys were more apt to explore different aspects of their environment than were the treated monkeys because the treated monkeys had facilitated performance under conditions in which there were irrelevant stimuli present.
I have never been terribly satisfied by that explanation; it is really a "Just so" explanation that you make up after the fact. I do not know why this is happening, but your data are not inconsistent with other monkey data on TCDD, so I think your results are really interesting.

Yoshikawa: Thank you for your interesting suggestion. I would like to think it over once again with that in mind.

Koibuchi: Thank you very much. Any other questions? Please.

Q: I have two questions.
One has to do with the method of bisphenol A exposure. I accept the fact that it is extremely difficult to keep up forced oral administration. You placed a lot of emphasis on natural exposure. Doses are generally administered in the drinking water or by mixing a certain concentration in the

drinking water in experiments such as this. Why didn't you use this method?
The other thing I would like to know is, I always have my doubts about behavior models, especially because behavior is affected to an extreme degree especially according to the stage of the sexual cycle in females when the sexual cycle is completed. Because the data was from eight weeks, I think the sexual cycle may have completed. Have you considered this?

Yoshikawa: In a certain sense, the dose for this experiment was quite high, and because bisphenol itself is hydrophobic, it will not dissolve in water. Also, when mixed with water or food, there is the problem of whether or not it can be completely processed. In laboratory experiments, we started pilot tests with rats, and in a certain sense, using the methodology in which I was in control myself, I decided to use the method of forcing them to drink the substance through a probe.

Another thing we were interested in was, more than males changing into females or female abnormalities, comparing the difference among specimens of the same sex. All data is therefore analyzed in that way. We assessed whether a difference might appear in cases in which there was no difference between the sexes. As for the points indicated a little while ago, there was a difference in term point for each specimen in the group according to dose.

Q: Your data was not a comparison of the sexes, so in that sense I think it was proper. But if for example you made a comparison among females, when behavior was affected, for example when taken randomly, if the sexual cycle for example were to be extended, I think that could be indirectly assessed.

Yoshikawa: That is possible. I think you must therefore be extra careful with females.

Koibuchi: Any other questions? If not, thank you very much, Dr. Yoshikawa.