

Scientific Review of Vaccine Safety Datalink Information

June 7-8, 2000

Simpsonwood Retreat Center

Norcross, Georgia

Dr. Orenstein:

My name is Walter Orenstein. I'm Director of the National Immunization Program at CDC and I want to thank all of you for coming here and taking time out of your very busy schedules to spend the next day and a half with us. Not only do we thank you for taking time out, but for taking the time out on such short notice, and also putting up with what I gather those of us who are townies here didn't realize, but apparently the biggest meeting in Atlanta which has taken up all the hotel space and all of the cars, so I think many of you have had to take taxis here. We appreciate you putting up with this, but at least we did arrange the weather nicely and you can look out occasionally and see some beautiful trees.

I think I am particularly impressed with the quality of expertise. We truly have been able to get at very short notice some of the most outstanding leaders in multiple fields. That will be important in interpreting the data.

We who work with vaccines take vaccine safety very seriously. Vaccines are generally given to healthy children and I think the public has, deservedly so, very high expectations for vaccine safety as well as the effectiveness of vaccination programs.

Those who don't know, initial concerns were raised last summer that mercury, as methylmercury in vaccines, might exceed safe levels. As a result of these concerns, CDC undertook, in collaboration with investigators in the Vaccine Safety Datalink, an effort to evaluate whether

there were any health risks from mercury in any of these vaccines.

Analysis to date raise some concerns of a possible dose-response effect of increasing levels of methylmercury in vaccines and certain neurologic diagnoses. Therefore, the purpose of this meeting is to have a careful scientific review of the data.

This is not a policy making meeting. Vaccine policy making will take place after this consultation as part of the Advisory Committee on Immunization Practices, or ACIP deliberations. For those who don't know, vaccine policy for CDC is really set through the recommendations of the Advisory Committee on Immunization Practices, or the ACIP. Thus, this is a scientific review to evaluate the quality of the scientific data. Our goal is to assure our policies are based on the best available scientific information.

This is what is called an individual simultaneous consultation. What that means is each consultant will be asked for their opinion publicly on questions which Roger Bernier will bring up in a few moments.

Although it will be of interest to see if the individual consultants tend to agree on particular issues, there is not the need to reach complete consensus. Your individual opinions should be very useful to the ACIP as it deliberates afterwards on policy options with regard to mercury in vaccines.

We hope you will participate in discussions, listen to the comments of others and form your own opinions during this day and a half meeting.

Again, we thank you very much for coming here and we look forward to a productive consultation.

In order to start, since many of us don't know each other, perhaps if we could go around the room and introduce each other. Let me ask maybe John, if you want to start.

Dr. Modlin: Certainly. I am John Modlin. I'm Chair of the ACIP and a member of the faculty at Dartmouth Medical School.

Dr. Stehr-Green: I'm Paul Stehr-Green. I'm an Epidemiologist by training. I am an Associate Professor of Epidemiology at the University of Washington School of Public Health and Community Medicine and I'm also a consulting Epidemiologist for the Northwest Portland Area Indian Health Board.

Dr. Stein: I am Marty Stein. I am on the faculty of Pediatrics at the University of California, San Diego where I am a General Pediatrician as well as Behavioral Pediatric and I co-chaired the American Academy of Pediatrics recent practice guideline on the diagnosis and evaluation for ADHD.

Dr. Saari: I'm Tom Saari, Professor of Pediatrics, University of Wisconsin in Madison and the Division of Infectious Diseases in Pediatrics. I'm also on the AAPCOID and I've represented liaison relationships to a number of national organizations.

Dr. Word: I'm Bonnie Word, I am at the State University of New York in Stony Brook. I am also a member of the ACIP.

Dr. Rennels: I'm Peggy Rennels, a pediatric infectious disease specialist at the Center of Vaccine Development, University of Maryland. I am a member of the ACIP and the AAP Committee on Infectious Diseases.

- Dr. Rapin: I'm Isabelle Rapin. I'm a Neurologist for children at Albert Einstein College of Medicine. I'm interested in developmental disorders, in particular language disorders and autism most recently.
- Dr. Sullivan: I'm Kevin Sullivan. I'm an Epidemiologist at Emory University, with the Department of Epidemiology and the Department of Pediatrics.
- Dr. Clarkson: I'm Tom Clarkson and I come from an area of frozen tundra in Rochester, New York. I've been associated with the mercury program through Rochester for a long time.
- Dr. Koller: Loren Koller, Pathologist, Immunotoxicologist, College of Veterinary Medicine, Oregon State University.
- Dr. Smith: I'm Natalie Smith, Director of the Immunization Program at the California State Health Department.
- Dr. Johnson: David Johnson. I'm the State Public Health Officer in Michigan and a member of ACIP.
- Dr. Clover: I'm Richard Clover, present chair of the Department of Family and Community Medicine, University of Louisville. I'm a member of the ACIP.
- Dr. DeStefano: I'm Frank DeStefano, Medical Epidemiologist in the National Immunization Program. I'm the project director of the Vaccine Safety Datalink.
- Dr. Chen: I'm Bob Chen, I'm Chief of Vaccine Safety and Development at the National Immunization Program at CDC.

- Dr. Davis: I'm Bob Davis. I'm one of the Associate Professors of Pediatrics and Epidemiology at the University of Washington. I am also one of the investigators.
- Dr. Johnston: I'm Dick Johnston, I'm an Immunologist and Pediatrician, now at the University of Colorado School of Medicine and National Jewish Center for Immunology and Respiratory Medicine. Adverse events related to vaccines has been of particular focus and interest for me mostly through serving on a series of committees dealing with the relationship between the vaccine and punitive adverse events.
- Dr. Bernier: I'm Roger Bernier, the Associate Director for Science in the National Immunization Program.
- Dr. Gerber: I'm Michael Gerber, I'm a medical officer at the National Institute of Allergy and Infectious Diseases of the National Institutes of Health. I'm also a member of the American Academy of Pediatrics Committee on Infectious Diseases.
- Dr. Mast: Eric Mast, I'm a Medical Epidemiologist with the Hepatitis Branch at CDC.
- Dr. Howe: Barbara Howe, I'm in charge of the clinical research group for vaccine development for Smith Kline Beecham in the U.S.
- Dr. Phillips: Bill Phillips from Seattle, Washington where I'm in private practice of Family Medicine. I'm here representing the American Academy of Family Physicians where I chair the Commission on Clinical Policies and Research.
- Dr. Caserta: Vito Caserta, I'm the Chief Medical Officer for the Vaccine Injury Compensation Program.

Ms. Ray: I'm Paula Ray, I'm with the Northern California Vaccine Study Center and I'm project manager for that site, for the VFC.

Mr. Lewis: I'm Ned Lewis. I'm the Data Manager at the Northern California Kaiser Vaccine Study Center.

Dr. Jones: I'm Dennis Jones. I'm a Toxicologist and Veterinarian. I'm the Assistant Director for Science, Division of Toxicology, ATSDR.

Dr. Egan: I'm Bill Egan, Acting Director for the Office of Vaccines Research and Review at FDA.

Dr. Deal: My name is Carolyn Deal. I'm the Acting Deputy Director of the Division of Bacterial Products at CBER at the FDA.

Dr. Pratt: I'm Douglas Pratt. I'm a Medical Officer in the Office of Vaccines at FDA.

Dr. Staub: I'm Ted Staub, I'm the Global Head of Biostatistics and Data Systems for Aventis Pasteur.

Dr. Sinks: My name is Tom Sinks. I'm the Associate Director for Science at the National Center for Environmental Health here at CDC and I'm also the Acting Division Director for the Division of Birth Defects, Developmental Disabilities and Disability Health.

Dr. Hadler: Steve Hadler, Medical Epidemiologist, National Immunization Program.

Dr. Mawle: I'm Alison Mawle, I'm the Vaccine Coordinator for the National Center for Infectious Diseases at CDC.

Dr. Rodewald: Lance Rodewald, I'm a Pediatrician and Associate Director for Science in the Immunization Services Division at CDC.

Dr. Kurz: Xavier Kurz, I'm Physician and Epidemiologist from Brussels, Belgium. I'm representing the European Agency for the Evaluation of Medicinal Products.

Dr. Pless: I'm Robert Pless, I'm a Medical Epidemiologist with the Vaccine Safety and Development Branch at the Immunization Program.

Dr. Clements: John Clements, the Expanded Program on Immunization, WHO, Geneva.

Mr. Schwartz: Ben Schwartz. I'm in the Epidemiology and Surveillance Division at NIP.

Dr. Myers: Martin Myers, I'm the Acting Director of the National Vaccine Program Office.

Dr. Guess: I'm Harry Guess. I'm head of the Epidemiology Department at Merck Research Labs.

Dr. Brent: I'm Robert Brent from Thomas Jefferson University and the Dupont Hospital for Children. I'm a Developmental Biologist and a Pediatrician.

Dr. Blum: I'm Mike Blum. I'm from Safety Surveillance and Epidemiology at Wyeth.

Dr. White: Good morning, I'm Jo White from North American Vaccine. I'm in charge of clinical development and research there.

Dr. Weil: I'm Bill Weil, an old Pediatrician who is representing the Committee on Environmental Health of the Academy at this moment.

Dr. Cordero: Good morning, Jose Cordero, Deputy Director of the National Immunization Program.

Dr. Chu: Susan Chu, Deputy Associate Director for Science, National Immunization Program.

Dr. Rhodes: Philip Rhodes, a Statistician in the National Immunization Program.

Dr. Verstraeten: I'm Tom Verstraeten, EIS Office at National Immunization Program.

Dr. Oakes: I'm David Oakes, the Chair of Biostatistics at the University of Rochester.

Ms. Heaps: I'm Wendy Heaps, Health Communications Specialist with NIP.

Dr. Orenstein: I'd like now to turn the meeting over to Roger Bernier who will give us a chronology of events, charge to consultants and talk about our Chairman and

Dr. Bernier: I believe the person has arrived with everyone's folders. I apologize that we didn't get them all here earlier this morning, but they should all be here now. You should each have a tent with your name on it and you should have a name badge. The information in there is just an agenda and a copy of the information that was handed out before the meeting.

I also want to reiterate a couple of points made by Walt Orenstein. Number one, that we have assembled quite an impressive array of expertise for this meeting. Some of you wondered why you were invited and worried that you wouldn't be able to provide good advice. We are not expecting any one person to be able to cover all of these

topics. As you can see, we have amassed quite an array of expertise, so we feel we have covered all the bases for the questions that will arise, but no one individual is expected to be able to comment on all of this.

The other thing I want to say is to reiterate the thanks of the CDC and the National Immunization Program. For some of you who have made yourselves available, you were not available when you were telephoned and invited, but some of you have been willing to change your schedules to make yourselves available and we genuinely appreciate that.

Dr. Bernier:

Let me talk just a little bit about the procedures today. I hope you have all received an agenda, but very quickly to give you an idea and feeling for how this day and the meeting has been planned to unfold, that isn't to say that is the way it is going to happen, but the critical presentation this morning is really the one by Tom Verstraeten, which is scheduled at eleven o'clock. We have some introductory presentations prior to that, but that is the critical presentation presenting the basic information. We have allowed an hour for discussion of that presentation. There is more discussion time at the end of the day if we need it, but we hope to get that presentation in with ample time for discussion before lunch.

Then Bob Davis will give a presentation about results of chart reviews which is supplemental to Tom. Then an independent review by Phil Rhodes. Then a comment on biologic plausibility and consistency by Dr. Koller, then we will have the break and ample time for discussion.

Tomorrow we begin with discussion for any residual questions, then we will get to the individual consultants opinions. You will be asked your opinions and we'll go through that in a minute as to what the questions will be.

The hope will be that you can look at those questions today and maybe prepare some notes so that you can talk from your notes tomorrow. Then we'll give you a clean sheet as you may want to revise what you wrote after you hear peoples opinions.

Then we will have a discussion about research and potential next steps at the end of the morning and then we will ask for your opinions again about what you think about any next steps.

Walt, do you have those opinion sheets? Okay, they are in the folders. There should be two of them. As I said, one that you might want to fill out this evening and take notes on, and then a clean one for any revisions that you may want to make after that.

There are about five or six different groups here. You may have figured that our from the introductions. I believe there are eleven consultants from CDC. There is also a list of participants. It has been distributed, so you should have a list of participants. There is a list of participants which identifies the eleven CDC consultants. They are the ones who will be asked to fill out these sheets. Others of you may want to do that. Feel free to do that, but you are not under an obligation to do it.

Another thing I would like to mention is that we will have a rapporteur, Dr. Paul Stehr-Green who is an Epidemiologist who introduced himself earlier. Paul is going to serve as rapporteur and we have allowed a half hour for a summary. I don't know if he will use it all, but he has a half hour to give a summary at the end of the second day tomorrow, so he may corral you now and then during the meeting to ask for clarification of some things or points that you have made. But you will know why because Paul has been asked to do that.

I briefly wanted to show people the immunization schedule.

Dr. Orenstein:

Can I make a very quick announcement? In addition to this being a simultaneous individual consultation on the part of the CDC, this is also going to be the initial meeting of the ACIP work group on Thimerosal and immunization, and the work group at the moment will consist of the five voting members of the Committee that are here. We will almost certainly expand the work group prior to the full ACIP meeting in about two weeks, but it will be important for the work group to get together at this meeting. I am hoping that the five of us can get together after dinner this evening. We will find a place and begin to discuss the various options and lay out the options for the full Committee in two weeks.

Dr. Bernier:

For some of you who don't work with vaccines every day, some of the consultants, just to let you know the focus of this. We are not likely to focus on all the vaccines today, but the three that are going to be of primary interest because they are given early in life include the Hepatitis B vaccine, which is recommended in three doses, and the DTP vaccine, diphtheria, tetanus, pertussis, which you will hear about and also haemophilus influenza type B which you see here according to this schedule. There will not be much discussion today about polio, measles, mumps, rubella, varicella or Hepatitis A. These vaccines have not, and do not contain Thimerosal. The focus is going to be about Hepatitis B, DTP and H. flu vaccine.

Now the other thing that I thought would be helpful is to try to provide a brief summary of the chronology of events that surround Thimerosal. This is not the first time some of us have heard of this preservative.

Basically there was a Congressional Action in 1997 requiring the FDA to review Mercury in drugs and biologics. In December 1998 the Food and Drug Administration had called for information from the manufacturers about mercury in their products.

There is a European group of regulation authorities and manufacturers that met in April of 1999 on this, who at that time noted the situation, but did not recommend any change.

In the U.S. there was a growing recognition that the cumulative exposure may exceed some of the guidelines. There are three sets of guidelines that are much in discussion. One from ATSDR, one from FDA and one from the Environmental Protection Agency. These guidelines are not all exactly the same. There was a recognition that the cumulative exposure that children receive from vaccination may actually exceed at least one of the guidelines that is recommended, that of the EPA. That caused a concern which resulted in a joint statement of the Public Health Service and the American Academy of Pediatrics in July of last year, which basically stated that as a long term goal, it was desirable to remove mercury from vaccines because it was a potentially preventable source of exposure. And if it was able to be removed, that it should be removed as soon as possible. That goal was agreed upon. In the meantime, there was postponement recommended for the Hepatitis B vaccine at birth. Also at that time, the FDA had sent a letter to manufacturers asking them to look at the situation with their products to see what could be accomplished as soon as possible.

There was a public workshop on Thimerosal in August of 1999. Dr. Myers will tell you a little bit about that this morning. In September of 1999, one of the Hepatitis B vaccines had removed Thimerosal from the product, so the

recommendation was made to resume use of Hepatitis B vaccine at birth.

Since that time, I believe in October of 1999 the ACIP looked this situation over again and did not express a preference for any of the vaccines that were Thimerosal free. They said the vaccines could be continued to be used, but reiterated the importance of the long term goal to try to remove Thimerosal as soon as possible.

Since then, I don't think there have been any major events. What has happened in the meantime is we have continued to look at this situation and that is what you are going to hear more about at this meeting.

Are there any questions about any of this?

Dr. Clarkson: Could we get copies of these transparencies that you are showing?

Dr. Bernier: Yes, we will arrange for that.

The next thing I would like to do, I have asked Dr. Dick Johnston to chair this meeting and he has been very gracious to accept that invitation from CDC. So at this point I would like to turn the meeting over to Dr. Johnston who will chair the meeting and keep us on track as much as possible. Thank you.

Dr. Johnston: Thank you, Roger. Jameka urged you to relax and it is my responsibility to be sure you don't relax during the presentations at least so that at the end, we who are consultants can vote with the greatest amount of understanding and knowledge of what the issues are. Otherwise, I am not going to take any more time. Marty, next?

Dr. Myers: In the meantime, I think Dixie Snyder has walked in. If anyone else has come in late that hasn't been introduced, you might want to introduce yourselves. Dixie, would you like to say hello?

Dr. Snyder: Good morning. I'm Dixie Snyder, the Associate Director for Science at CDC and the Executive Secretary for the ACIP.

Dr. Walker: I'm Alex Walker, I'm Chair of the Epidemiology Department at the Harvard School of Public Health.

Dr. Johnston: A lot of you were at the conference that I'm going to summarize, so if I omit something or over interpret something, please jump in.

The conference that Roger was alluding to was a hot, sultry couple of days at Bethesda at the Lister Auditorium last August where the National Vaccine Advisory Committee and the Inter-Agency Working Group on Vaccines convened a special meeting to consider Thimerosal in vaccines. Obviously a pertinent topic for this morning.

I think one of the major take home lessons was that we should have had that meeting in advance of many of the public health decisions that were made last summer, although that wasn't possible, but it would have been desirable to have a meeting such as we are having today to consider the data first.

Thimerosal is in many vaccines because it is a preservative and it lowers the rate of bacterial and fungal contamination that may occur during the manufacturing process, packaging and the use of vaccines in the field. It is particularly a concern in multi-dose vials because of the issue of re-entry multiple times in the vials, and it is also important in the manufacturing process for a number of

vaccines including inactivated influenza and some of the earlier DPT vaccines, and is a constituent of all DPT vaccines, but not all DTAP vaccines.

There are three licensed preservatives in the United States, Thimerosal, ethynyl and phenol. We won't talk about the other two today, but I thought I should mention them. Thimerosal is the most active and it has been utilized in vaccines since the 1930's.

At the time of the meeting last summer, there was only one licensed product containing hepatitis B vaccine that did not contain Thimerosal. That was a combination vaccine with HIV that was intended for use in two months or older, so the issue was that all of the vaccines available for the birth dose contained Thimerosal. In addition, many of the DTAP vaccines and the HIV vaccines, many, but not all, contained Thimerosal.

Thimerosal functions as an anti-microbial after it is cleaved into ethylmercury and thiosalicylate, which is inactive. It is the ethylmercury which is bactericidal at acidic PH and fungistatic at neutral and alkaline PH. It has no activity against spore forming organisms.

There is a very limited pharmacokinetic data concerning ethylmercury. There is very limited data on its blood levels. There is no data on its excretion. It is recognized to both cross placenta and the blood-brain barrier.

The data on its toxicity, ethylmercury, is sparse. It is primarily recognized as a cause of hypersensitivity. Acutely it can cause neurologic and renal toxicity, including death, from overdose.

Because of the limited data for ethylmercury and its physical chemical similarities to methylmercury, it was the

consensus of the meeting that in the absence of other data, that chronic exposure to methylmercury would need to be used to assess any potential neurodevelopmental risks of ethylmercury, although it was recognized that we needed data specifically on ethylmercury.

We learned a great deal about the toxicity of ethylmercury from animal studies, accidental environmental exposures, and studies of island populations who consume large amounts of predator fish that contain high concentrations of ethylmercury.

We learned that ethylmercury is ubiquitous and that assessments of exposure by infants would need to include environmental exposures, maternal foods, whether the baby was nursed or not, as well as their exposure to vaccines.

Specialists in environmental health have extrapolated from those types of studies to establish safe exposure levels, and this is an important emphasis I would like to make on chronic, daily exposure to ethylmercury that incorporate wide margins. That is three to ten fold to account for data uncertainties.

As an aside, we found a cultural difference between vaccinologists and environmental health people in that many of us in the vaccine arena have never thought about uncertainty factors before. We tend to be relatively concrete in our thinking. Probably one of the big cultural events in that meeting, at least for me, was when Dr. Clarkson repetitively pointed out to us that we just didn't get it about uncertainty, and he was actually quite right. It took us a couple of days to understand the factor of uncertainty in assessing environmental exposure, particularly to metals.

If methylmercury were applied as a surrogate for ethylmercury, then some combinations of vaccines, according to the recommendation that Roger showed us, could result in some children having organomercurial exposure that exceeded some of those guidelines. Specifically the EPA guideline.

There were a number of things that we got a consensus on in that meeting. First is that there was no evidence of a problem, only a theoretical concern that young infants' developing brains were being exposed to an organomercurial.

We agreed that while there was no evidence of a problem, the increasing number of vaccine injections given to infants was increasing the theoretical mercury exposure risk.

We agreed that the greatest risk for mercury exposure from vaccines would be to low birth weight infants and to infants born prematurely.

We agreed that it would be desirable to remove mercury from U.S. licensed vaccines, but we did not agree that this was a universal recommendation that we would make because of the issue concerning preservatives for delivering vaccines to other countries, particularly developing countries, in the absence of hard data that implied that there was in fact a problem.

There were a lot of uncertainties that we left the meeting listing. The first was chronic versus episodic exposure, oral versus parental exposure, ethyl versus methylmercury, the dose of mercury on a per kilogram base at birth and subsequently the issue of pre-term versus term birth.

We did then discuss both theoretical and real disease burden risk. We saw some compelling data that delaying

the birth dose of hepatitis B vaccine would lead to significant disease burden as a consequence of missed opportunity to immunize.

We have since seen those initial recommendations in July a year ago, a reduction in appropriate use of hepatitis B immunoprophylaxis to infants born to mothers who were hepatitis B surface antigen positive.

Dr. Clarkson made the compelling point that delaying the birth dose from day one or two until two or six months later would have a limited impact on the cumulative mercury exposure, and the point was made that the potential impact on countries that have 10% to 15% newborn hepatitis B exposure risk was very distressing to consider.

We concluded the meeting with a research agenda, and as that is on the agenda for tomorrow, Alison Mawle and Mike Gerber were on that panel so they can contribute specifically.

A couple of issues that were raised and probably are worth raising in the context of what we are going to discuss in this consultation, what contribution does vaccine mercury play in the isolated communities where mercury exposure was examined very carefully? What are the pharmacokinetics of excretion of ethylmercury? And then at the end of the meeting ironically, Walt Orenstein asked the most provocative question which induced a great deal of discussion. That was, should we try to seek neurodevelopmental outcomes for children exposed to varying doses of mercury by utilizing the Vaccine Safety Datalink data from one or more sites?

The discussion that followed that, and I did review the transcripts of this in preparation, is very interesting. Drs. Gerber and Clarkson especially, but a number of others of

us also, expressed grave concerns that the many confounding co-variables would make such data very difficult to evaluate. Dr. Halsey made a very impassioned plea that we do carefully controlled studies to in fact address the issues specifically, and that such studies be conducted by neurodevelopmentalists and environmental scientists employing specific endpoints of their study.

I suspect that today we will consider many of those confounding variables from the Vaccine Safety Datalink.

Finally I would like to mention one more issue. As you know, the National Vaccine Program Office has sponsored two conferences on metals and vaccines. I have just recounted a summary of the mercury, the Thimerosal in vaccines. We just recently had another meeting that some of you were able to attend dealing with aluminum in vaccines. I would like to just say one or two words about that before I conclude.

We learned at that meeting a number of important things about aluminum, and I think they also are important in our considerations today. First, aluminum salts, and there are a number of different salts that are utilized, reduce the amount of antigen and the number of injections required for primary immunization.

Secondly, they don't have much role in recall immunization, but it would represent a significant burden to try and develop different vaccines for primary and subsequent immunizations.

Aluminum salts are important in the formulating process of vaccines, both in antigen stabilization and absorption of endotoxin.

Aluminum salts have a very wide margin of safety. Aluminum and mercury are often simultaneously administered to infants, both at the same site and at different sites.

However, we also learned that there is absolutely no data, including animal data, about the potential for synergy, additivity or antagonism, all of which can occur in binary metal mixtures that relate and allow us to draw any conclusions from the simultaneous exposure to these two salts in vaccines.

Thank you very much.

Dr. Johnson: Marty, the ethylmercury has been painted with the methylmercury brush, and maybe we will discuss this later, but are they metabolized equivalently, exactly equivalently, partially, differently?

Dr. Myers: I'm not sure that I'm confident to answer that. Dr. Clarkson, if I recall, when asked that question specifically at the mercury conference said that we should assume that their excretion was similar, but that might well not be the case. That would be the worst case scenario.

Dr. Johnson: Well, we have a discussion tomorrow on biologic plausibility and maybe that will deal with that. Dr. Clarkson was quoted as saying that delaying HepB for six months or so would not affect the mercury burden, but I would have thought that the difference was in the timing. That is you are protecting the first six months of the developing central nervous system. Is that not?

Dr. Myers: I probably should allow him to speak for himself, but my interpretation was that the health guidelines were established based on a chronic, every single day exposure, and that a single day exposure, if I can quote him

accurately, wouldn't change the blood levels one femtogram.

Dr. Clarkson: I'm not sure where this statement came from, but I'm glad you raised it. Since the dose is the same for each vaccine administration, clearly there is a body weight difference after six months, so the actual dose per tissue will be lower at six months. I'm not quite sure what the question was.

Dr. Johnson: Well, maybe Isabelle should do this. I don't want to spend too much time, but the time of exposure, that is the central nervous system of a newborn and so forth, does that make a difference in the biologic plausibility related to central nervous system effects that are under consideration?

Dr. Clarkson: It could make a difference certainly. The guidelines that Dr. Myers is talking about is based on prenatal exposures and perinatal exposures. As far as I know the literature, there just isn't that much evidence one way or the other as to whether exposure shortly after birth or exposure at six months would make a difference. In theory it could, but I don't know of any studies that have actually tested that.

There is an issue that the pharmacokinetics might be different, too. Again, this is all animal work, but the animal studies suggested, for example, a suckling animal does not eliminate methylmercury until the end of the suckling period, and there is a mechanism on the study for that. So this is not known for humans. So there could be an age difference in the excretion rates.

Dr. Rapin: I am not an expert on mercury in infancy. The diseases that neurologists know about mercury in infancy have more to do with the peripheral nervous system than with the central nervous system. I know of at least one child that was exposed to mercury and developed a very severe

neuropathy, but I don't know whether the child, if one would test her carefully, had any cognitive deficits.

I don't know if anyone has looked at the literature of the old Pinks disease which was present in the twenties or thirties when mothers wore shields that contained mercury. I really don't know, so I'm sorry.

Dr. Snyder:

I think the issue at the meeting that I thought Dr. Clarkson was telling us was that we were focused on the amount of mercury in a particular dose of vaccine, and we needed to think beyond that in terms of what that meant for blood levels and therefore tissue levels, and then specifically the target organs. If we look at that single dose, let's say of hepatitis B vaccine, that single dose was not going to ratchet up the blood level. Whatever it was, for background reasons from food intake of the infant or the mother, that one dose was not going to make a major change in blood levels, and therefore major changes in tissue levels. That's the way I interpreted the statement at the meeting. Which is not to say it's unimportant, but it was a small amount relative to all the other intake.

Dr. Clarkson:

As you know, there is a paper just published on this now which I guess many of you in pediatrics have a copy of now. That's right, if you are given mercury day by day as the guidelines are based on, whether it's EPA, ATSDR or FDA, these are based on a constant daily exposure and at least for adults it would take almost a year to get to study state levels. Whereas we are just considering one single dose for vaccines.

But nevertheless, a single dose from vaccines can raise blood levels by a certain amount. We now have one paper showing that in fact it does and the level it is raised to is reasonable. It's reasonable for what we would expect the dose to be and what the body weight should be, and these

of course are in very low birth weight infants that the report was on.

Dr. Brent:

It's just the sensitivity of the central nervous system, based on the mechanism that's involved in producing the end result. You know the thalidomide data taught us that autism is related to the high brain and it produces it in the 22nd day of gestation, while the central nervous system from the standpoint of mental retardation, its most sensitive period is in the eighth week to the fifteenth week. That's when we see the neuro-maturation.

You are talking about miolimitation. I don't know of any data of whether there is a sensitive period miolionization or if you have a high enough dose you can affect miolionization throughout the period of miolionization.

I think you have to recognize that each of the developmental problems that have been evaluated here have a different stage where they are most sensitive from environmental factors.

Dr. Johnson:

Are any of them different from birth, term birth to six months?

Dr. Brent:

In Hiroshima, Nagasaki, you had severe mental retardation after 75 rads. If you give 75 rads to an infant, nothing will happen with regards to their central nervous system development. So you have this changing sensitivity throughout embryogenesis and early childhood development that makes it very difficult to generalize.

Dr. Johnson:

So the answer is that we don't know. Between birth and six months there is no reason particularly, based on data at least, to be concerned that shifting the exposures back toward birth is any more risky than waiting till six months.

Dr. Myers: The one thing that was a take away from that meeting was that if there were an increased risk, it would be in the low birth rate and preterm infants.

Dr. Sinks: I wanted to ask an unrelated question, and this has to do with potentially looking at confounding as we go through this. You mentioned the issue of aluminum salts. I know it's an issue, but I don't know the specifics of it. I wonder is there a particular health outcome that has been of concern that is related to the aluminum salts that may have anything to do with what we are looking at here today?

Dr. Myers: No, I don't believe there are any particular health concern that was raised. It was raised as an issue, and clearly it's a confounding issue in that exposure to vaccine includes exposure to things other than Thimerosal.

Dr. Weil: Two things. One, up until this last discussion we have been talking about chronic exposure. I think it's clear to me anyway that we are talking about a problem that is probably more related to bolus acute exposures, and we also need to know that the migration problem and some of the other developmental problems in the central nervous system go on for quite a period after birth. But from all the other studies of other toxic substances, the earlier you work with the central nervous system, the more likely you are to run into a sensitive period for one of these effects, so that moving from one month or one day of birth to six months of birth changes enormously the potential for toxicity. There are just a host of neurodevelopmental data that would suggest that we've got a serious problem. The earlier we go, the more serious the problem.

The second point I could make is that in relationship to aluminum, being a nephrologist for a long time, the potential for aluminum and central nervous system toxicity

was well established by dialysis data. To think there isn't some possible problem here is unreal.

Dr. Johnson: Thank you, Bill, for your comments. As an old pediatrician, I had that same kind of feeling. That there must be a difference with age.

Dr. Myers: Just to not leave that as a hanger though, our metal experts, and we had quite a collection of people. We held the aluminum meeting in conjunction with the metal lions in biology and medicine meeting, we were quick to point out that in the absence of data we didn't know about additive or inhibitory activities. We should not conclude necessarily that they would be additive. I think that was Tom's point, in the absence of a health endpoint, we needed to be very careful. But I did want to raise the issue because it was a major issue of discussion there, that we did have binary salt exposure and we probably needed to understand more about that.

Dr. Johnson: Thank you very much, Marty. I'm sure we'll hear more from Dr. Koller this afternoon.

Frank DeStefano is going to introduce us to the Vaccine Safety Datalink Study.

Dr. DeStefano: The analyses you will be discussing for most of the morning come from the Vaccine Safety Datalink. I'm going to give you a quick overview of what the project is and then some of the data.

This is a project collaboration between the CDC's National Immunization Program and four large health maintenance organizations listed here, Group Health Cooperative in Seattle, Northwest Kaiser in Portland and Northern and Southern California Kaiser. They have a current enrolled population between them of over 6 million people.

For a little history on the project, it was begun to have a large population to address primarily rare potential vaccine safety problems. It began in March of 1991 at three sites, then the Southern California site began contributing data in October of 1992.

The size population of between zero and six years old, this will be cumulative, over the nearly ten years of the project, I think we're probably over 2 million children now.

The concerns about HMOs sometimes have to do with their representativeness, at least in terms of data that the HMOs have been able to compare with the areas that they serve, they tend to be fairly similar in terms of ratio, ethnic, characteristics, age and such. Then we have expanded to include adolescents and adults, but we won't be discussing those today.

So the Datalink, this is sort of a schematic of what we are talking about here. The study begins with computerized data that the HMOs collect primarily for administrative and medical care purposes. They are collected for different reasons. The goal of the Datalink is to try to use those data and combine them to do vaccine safety epidemiologic studies.

There are three main types of data that we use. Automated vaccination records. These are computerized immunization tracking systems if you will. Some of these could be considered the prototype vaccination registries. Obviously that's a key for doing vaccine safety studies. The other main source of data deals with identifying health outcomes. I will talk more about those in a subsequent overhead. Then another key component is getting information on patient care characteristics, such as date of birth, gender and particularly important are dates when members enroll

and disenroll and the population which is critical for keeping track of the population under observation. Those data are sent to us at CDC. Each of the members of the HMO have a unique identifying number that is used to link among the various data sources.

We at CDC serve sort of as a data coordinating center. We combine the data from all the four HMOs and do some of the combined data preparations and some of the analyses which we will be talking about.

As I mentioned, these are computerized data bases. I am sure you are all familiar with the potential concerns and limitations of computerized data bases. We have done quite a bit of validation, particularly of the vaccine data, and this is some results from three of the HMOs about the vaccines that are going to be of primary interest in today's study.

In Northern California Kaiser NCK, you can see what their sensitivity and positive predictive value is. What sensitivity means here is that if a vaccination was in a hard copy medical record or log, it was actually captured by the automated data system. This was done by doing some actual chart extractions and comparing what is in the computerized data base with the hard copy records. So for DTP, 98% of the time, if it was in a hard copy record it was captured in the computer data base in Northern California.

The positive predictive value means if it is identified in the computerized data base, when you go to the hard copy record it's there. Basically you can see these results. For Northern California there is very high agreement for all the vaccines of interest.

For Group Health, which is going to be the other main HMO that contributes to these analyses, you see the

agreement is fairly high, although not quite as high. Hepatitis B tends to be a bit low, and I think that primarily is because of capturing the birth dose. The hospital's HMO birth dose some times didn't tend to get into the data bases as well in the early years. I think Tom in his presentation will probably have some more to say about this.

This is primarily where we determine outcomes. The primary sources are hospital discharge diagnoses, and all HMOs have these. Then there are treatment records from clinics. For the conditions that we are talking about today, most of these are treated primarily in outpatient clinics and not all the HMOs had outpatient records. They were electronic records that they provided to us. It was most complete for Northern California and Group Health, so that is why you will see the analyses are restricted predominantly to those two HMOs.

We also have emergency room visits and can get Death Certificate autopsy reports, and if need be we have a variety of ancillary data sources, but we did not use any of those for these current analyses.

The sort of prototypical analytical approach is to use these computerized data. Here is a screening. Usually because of the problems with automated data in terms of the validity of the diagnoses, et cetera, the computerized data we use usually as a screening analysis. Primarily to see if there is any preliminary assessment of vaccine outcome associations, or sometimes it is used as a way to identify possible cases of a condition. Usually we go to a next step. We have found it is necessary to go to a next step for more detailed analysis, and usually this involves chart reviews. Some times it actually involves interviews of parents or patients. These more detailed chart reviews are necessary to validate the outcomes of interest to make sure what those computerized codes actually represent in terms of what was

written in the medical charts, or in terms if you come up with a more standardized case definition. Also to confirm when the date of occurrence was, or some times to get a more reliable onset or incidence date and to verify vaccination history. We tend to be more comfortable using the automated vaccination histories.

Then importantly to get additional risk factors, confounding information or other information on clinical details. Basically what we have in the computerized data on risk factors is gender and date of birth in essence.

I will just give a little background on how this analysis on Thimerosal developed. I guess after the August meeting in Bethesda that Marty has told you about, a Thimerosal working group was convened. Michael Gerber coordinated that working group and it included representatives from several public health service agencies, as well as people in academia and other organizations. Sort of an informal working group. As I understood it, the primary purpose seemed to be to come up with ideas for research to see if there was anything really to these theoretical concerns that had been raised about Thimerosal exposure. There were proposals about studies to look at what happens to body burden after vaccination. Michael may have some information on that later in the meeting.

One of the proposals that was made was to do a study that we will be talking about today. Looking at using the Vaccine Safety Datalink Project to look and see if there was any association between Thimerosal exposures as estimated through vaccinations received and selected outcomes. We weren't made aware of the concerns that had been raised at the August meeting.

At first it took a while for some people to understand the concept in the Datalink. I think by the second conference

call this concept got quite a bit of support, and we were encouraged to develop a protocol and such, which we did. We developed a protocol in conjunction with some input from this working group, as well as the Vaccine Safety Datalink investigators.

Basically the protocol called for a two phase study. The first stage was the screening of automated data for possible associations, and I want to emphasize this is what we will be talking about today. This was like a screening analysis. We did narrow down the conditions we were looking for to conditions that had been suggested to be related primarily to methylmercury, and those were primary neurologic, neurodevelopmental and renal outcome. Still within those there was a broad category of possible specific conditions and we didn't know if any of them would really have an association. So the idea was to do this screening analysis of automated data to see if there was any hint of association with any specific conditions, then the thought would be if anything came out we would go to the next step and do a confirmatory study or hypothesis testing study.

At the time we were thinking this would have been the usual chart review case control study. Since then in looking at the conditions that have seem to have come out as possible associations, I think we might rethink that strategy and hopefully we will have a chance to discuss what that phase two might be tomorrow morning.

Dr. Johnson: Thank you, Frank. Why don't we go right into Tom's presentation.

Dr. Gerber: Why did you choose these two of the four HMOs?

Dr DeStefano: They were the ones that had outpatient data.

Dr. Gerber: So the other two didn't have outpatient data?

Dr. DeStefano: Right.

Dr. Verstraeten: Good morning. It is sort of interesting that when I first came to the CDC as a NIS officer a year ago only, I didn't really know what I wanted to do, but one of the things I knew I didn't want to do was studies that had to do with toxicology or environmental health. Because I thought it was too much confounding and it's very hard to prove anything in those studies. Now it turns out that other people although thought that this study was not the right thing to do, so what I will present to you is the study that nobody thought we should do.

If I can have the next slide. Frank already mentioned to two phases that we originally considered for this study.

The design of the first phase, the screening phase, were we were looking for signals was as follows. We set it up as a cohort study using this automated VSD data. The exposure was to be mercury from Thimerosal containing childhood vaccines assessed at different ages of the children. The outcome was a range of plausible, neurologic and renal disorders. As plausible as I could find from the literature, anything that I could not exclude among the neurologic or renal disorders to be connected to mercury.

On the study of population, we selected children born between 1992 and 1997. We started in 1992 because we saw that is when the data became complete for the different HMOs.

They had to be born into two HMOs. We have already talked about that. The next condition was for these children to be continuously enrolled during the first year of life. We wanted to make sure that we captured all the vaccines given in the first year of life.

Finally, we excluded children that didn't receive at least two polio vaccines before the end of the first year of life. The idea here was that there is still children that are enrolled in the HMO, but may not be using the immunization facilities at the HMO. We thought that the polio would be the vaccine with the least contraindications, and two polio vaccines is what is routinely recommended, so we would exclude those children that had less than two.

The little asterisk indicates that this last condition was not in the original protocol. We added it as we started discussing our first findings.

There were some other children that we excluded. First, premature children. From the very start we said we were going to look at these children separately and there are specific reasons to do that. We know that premature children are not vaccinated in the same way as term babies. At the same time they are at higher risk for the outcomes, so we wanted to look at them separately.

Hepatitis B immunoglobulins. I think that is pretty obvious. Those would be vaccinated for hepatitis B and would have a higher likelihood of the outcomes.

Finally, we excluded children with congenital or severe perinatal disorders. That was also a condition that we added. It was not in the original protocol. The idea was to get as pure a group of children as possible. Children that we knew didn't have any problems before or at birth. I will come back to discuss this group later on.

The exposure we assessed cumulatively. We kept on counting the cumulative amount of mercury at different ages of the children. We calculated using these individual automated vaccination records and we assessed it at one, two, three and six months of age. We figured that the

earliest month of life would be the most sensitive to mercury, so we wanted to see what was going on during those months.

Then after we calculated that, we categorized these exposures by levels of 12.5. 12.5 is the minimum amount that any Thimerosal-containing vaccine has and namely hepatitis B has 12.5 micrograms of ethylmercury.

There is an assumption there that for the Hep vaccines, we weren't sure of this beforehand, but we confirmed later on that Hep vaccines in our cohort all contained Thimerosal.

Now for the outcomes, we looked at the neurologic and renal outcomes and we classified them into major categories. One of those is neurologic developmental disorders. Sometimes we refer to this category as NDD.

In this category we have all the outcomes that received any of these codes, which are on this slide. I will not go over all of them, but they include such things as autism, stammering and Tics. The largest group in here is under 315. That includes such things as speech and language disorders and coordination disorders. There is a very small group of mental retardation.

Another category were all the renal disorders which we put altogether into one large category. That goes from glomerulonephritis, nephrotic Syndrome and to renal failure. The major single code being used here is unfortunately the one that is called unspecified kidney and ureter disease.

Besides these two categories, we looked at some other neurologic disorders. Some of them we categorized in a group we called degenerative neurologic disorders, and then there was a final category of other neurologic

disorders, which we thought we could not put into one of the categories. It includes such things as epilepsy.

For our statistical analysis we used proportional hazard models. These models were stratified over the two HMOs, year and month of birth. Originally we had only thought of the year of birth, but very early on some people commented that was not specific enough. That we should add the month of birth. So we should compare children that were born in the same HMO, the same month and the same year so our cases and controls would come from within such a strain. Then we adjusted our analysis for gender. That is the only covariant we adjusted.

For each of these disorders I've mentioned before, we did a separate analysis. If we found within the cohort at least 50 cases, which was a very rough sample size estimate to detect, a relative risk of 2, so we said any disorder for which we find at least 50 cases we will do a separate analysis. All the other disorders we will just include in the overall category, but we will not look at them separately.

Now turning to the results. These are the number of children that we found. First of all, born in any of the two HMOs in that time period, we found a little more than 200,000 children.

This condition of being continuously enrolled eliminated quite a large number of those and we were left with 140,000. There was only a few thousand that didn't get their two polio vaccines by one year. There was about 5% premature children. There were very few children that received hepatitis b immunoglobulins and finally there was quite a large group, about 25%, that we excluded because of congenital or perinatal disorders. So we were finally left with about 110,000 children in our cohort.

Turning to the exposure, this is the different vaccines that contribute to the exposure at three months of age. I will focus a lot at this exposure of three months of age for reasons I will show you later on, mostly because it has the nicest distribution. At three months of age, children have from zero to over 75 micrograms of ethylmercury exposure from Thimerosal-containing vaccines. Zero, that's pretty obvious. They didn't get any vaccines. That's another important point to keep in mind. None of the vaccines, except for polio which is usually given together with DTP or haemophilus influenza, was Thimerosal free in our cohort. That means if the children don't have Thimerosal, it means they didn't receive any vaccines. Whether it's one, two, three or six months of age.

The next category would be the children that received one hepatitis B. One up from there would be the children that received two hepatitis B vaccines and no DTP or no Hib, which is haemophilus influenza. Or there is another possibility. There is DTP and HIB exist in a combination. It's called Tetramune. This vaccine contains 25 micrograms of ethylmercury, so it's only half of what the children get than when they get DTP and HIB separate, they get 50. If they get those two combined, they get 25.

The next category would be the same combination plus one hepatitis B.

Now at fifty, there is another two possibilities. Children can have received two hepatitis B vaccines before three months, and this combination vaccine, or no hepatitis B and the DTP and HIB separate, which I mentioned is 25 each. So that would add up to 50 also.

This combination vaccine was used only in one HMO, at Northern California Kaiser. In Group Health they don't use it. In Northern California Kaiser, the large majority of

children received the combination vaccine. That's why most of the children at Northern California Kaiser has a much larger contribution to this cohort than Group Health.

Finally the two top categories are both had one DTP and one HIB separately, combined with one hepatitis B or two hepatitis B. There are very few children that get more than 75 at three months. That would occur if they get more than one DTP or more than one Hib, together with two hepatitis B vaccines.

I know this slide is a bit busy, but if we take our time I think it will make sense. It's the distribution of ethylmercury from Thimerosal-containing vaccines at one, two, three and six months of age. This first part of the slide with the small numbers is the distribution at one month of age. Basically the distribution at one month is whether or not the child received hepatitis B or not. If they didn't receive hepatitis B, there was no mercury. If they received it, it was 12.5.

There is a few children who received their first DTP or their first HIB before they finished the first month of life, which I cut off as 31 days. So basically at the first month it is a dichotomous variable.

Going to two months, the distribution is quite similar. There are a few children who already received their DTP and HIB and possibly a second hepatitis B, but still the largest majority is in these categories.

At three months of age we get what resembles most normal distribution. Those are the categories which I've discussed with you on the previous slide, whereby the largest group is anywhere from 37.5, 50 or 62.5 micrograms of mercury. There is a few children in these low categories and there are very few children above 75.

At six months, the distribution becomes multi-modal with severe peaks at different levels of mercury.

Now what happens with our exposure over time. We'll probably talk a lot about temporal trends. This is the average mercury exposure at different months of age for the entire VSD cohort. Not just our cohort, but that includes two other HMOs. What happens over time is that between 1991 and 1992 there is a raise at all levels, which is due to hepatitis B. In 1991 hepatitis B was not much used in newborns. It was introduced mostly in 1992 and in some HMOs a bit later, and that's why we have this increase.

We have a small decrease after that which is mostly due to the introduction of this DTP-HIB combination vaccine, which reduces their cumulative mercury level.

In the end we have a slight increase again which happens when DTaP, the acellular DTP vaccine has introduced. That one did not exist in combination, or that was used very little. That made some of the HMOs go back to giving DTP and HIB separately, and that increased the levels of mercury again. There are some other factors that play a role, but most those changes would be due to those policy changes.

However, if we look at HMO by HMO, those trends are not as stable as they look for the entire VSD. This is for Group Health. This is only looking at three months of age. Those categories where I have lumped together, those categories below are equal to 25 micrograms. The one that jumps out mostly is the highest category, equal to or higher than 75 micrograms. What we will also notice here is that at Group Health, the exposure is higher than the other HMO, Northern California Kaiser. So at Group Health there is a jump from '92 to '93 for the highest category, and after that

it doesn't change much. It goes back down again in 1997. Basically the purpose of this slide is to show that those exposures are not entirely stable over the different years.

This is the same for Northern California Kaiser. What we see is that the most prevalent categories here are the ones of 37 and 50 micrograms at three months of age, indicating that the level of exposure at Northern California Kaiser is not as high as Group Health. Also the trends do vary a bit over time.

Turning to the outcomes, here are some crude data of the outcomes. The first slide shows you the total numbers for some of the outcomes by year of birth of the children. This is the total number of year of birth of the children, which is rather stable.

These are the numbers for the entire category of neurologic developmental disorders, where we see that it is basically the children born in the first years of our cohort, '92 to '95, who are contributing mostly to our outcomes, which is not surprising because the other children are just not old enough to be diagnosed with any of these disorders.

For speech, that's the same.

This is attention deficit disorder, which is another outcome on which we shall focus quite a bit. It's the same trend. It's mostly the children born in the earliest years. It's even more so for ADD where the children have to be older to be diagnosed. So it's good to bear in mind for some of the outcomes, we're talking mostly children born in the earliest years of our cohort.

This slide gives you the crude rates of the outcomes by level of exposure at three months of age. So what we have here are the categories of exposure at three months of age.

The numbers of each category which I showed you which vaccines are in each category, and now how many cases did we find in each category, followed by the rates which is taking in account person time, so these are rates by 1,000 person years.

I'll leave it up to you to think whether there are trends, yes or no. What is important to notice here is we have combined Group Health and NCK on this slide. On the next slide I will show you what happens when we separate them out.

Again, this is the entire category of neurologic developmental disorders. This is speech delay. This is attention deficit disorder. None of these numbers are in the text that you have received before coming here.

The purpose, many times we have been asked to provide the raw data to have a sense of what is going on and which numbers we are talking about. We are not looking in much detail at this time at these different rates.

For ADD, these three categories are lumped because the numbers become quite sparse.

This is what happens when we separate the two HMOs. What is important to notice is first of all, the overall incidence rates between the two HMOs differ substantially for some of the outcomes. We have a much higher rate of speed at Group Health compared to Northern California Kaiser. For attention deficit disorder, that number is not as high.

Secondly, the rates year by year or any trends that you might think one way or the other way can be quite different between the two HMOs. That is true for both disorders we have selected. One of the reasons we keep selecting these

disorders is because they have the most cases, so we avoid getting sparse results and we think some of the findings are significant for these disorders.

In summary, what we wanted to say about the data that we've shown you is the exposure varies quite a bit by HMO and over time. Secondly, the outcomes or the incidence of the outcomes also varies by HMO and time. Therefore, we think it is quite difficult to interpret crude results. If we come up with basic 2 x 2 tables, there would be a lot of confounding that we don't take into account. Therefore, we think we have to account for these different trends and differences by HMO in whichever risk analysis we do.

Now turning to the results of our proportional hazard models, we have compared in total 17 individual out of the 38 plausible outcomes. Meaning that 17 of those had at least 50 cases. Three of the grouped ones also had that number. We've compared those outcomes to seven different measures of exposure. The seven measures of exposure are in the text. They are the continuous measure at one, two, three and six months of age. Those are four, then there is the categorized exposure at three months of age. Finally, we have also included the dichotomized exposure at one and three months using the EPA limits as a cut off to difference between high or low exposure. That gives us seven measures of exposure.

From those risk analysis, excluding those dichotomized for EPA, we have found statistically significant relationships between the exposure and the outcome for these different exposures and outcomes. First, for two months of age, an unspecified developmental delay which has its own specific ICD9 code.

Exposure at three months of age, Tics. Exposure at six months of age, an attention deficit disorder. Exposure at one, three and six months of age, language and speech delays which are two separate ICD9 codes. Exposure at one, three and six months of age, the entire category of neurodevelopmental delays, which includes all of these plus a number of other disorders.

Now going into detail of some of these. The slides I will show now, they were also all in the original text which you have received. The results of the risk calculations for the exposure at three months of age are categorized into seven categories by 12.5 micrograms, and the last one is any exposure about 62.5 which is basically 75 micrograms.

The reference category in this calculation is the zero microgram category. In other words, the children that didn't receive vaccines.

For each of these categories what is shown is a point estimate and a 95% confidence intervals. Then these point estimates are linked by a continuous line to visualize a potential trend.

For each category I have shown here the number of cases for each category. Finally, this is a test for trend of these findings, which I have done by taking the exposure as a continuous variable. It gives you the 95% confidence intervals and the P value for the finding.

For the overall category of neurologic developmental disorders, the point estimates of the categorized estimates suggest potential trends, and the test for trends is also statistically significant above one, with a P value below 0.01. The way to interpret this point estimate which seems very low is as follows. That's an increase of .7% for each additional microgram of ethylmercury. For an example, if

we would go from zero to 50 micrograms of ethylmercury, we would have to multiple these estimate by 50, so that would give us an additional increase of about 35%, which is pretty close to the point estimate for this category. Or for the overall, we would have to multiple 75 micrograms to .7 and that would give us about one and a half for the relative risk.

If anyone has questions on this graph I will take them now because the next slides have similar slides and I think it is important to understand what these graphs represent.

Dr. Oakes: I take it you are only counting out after three months then?

Dr. Verstraeten: Absolutely.

Dr. Sinks: If I remember your exposure distribution, they were increasing not in actual micrograms, but in clumps because of the way the dose is applied. I wonder if it's appropriate to follow this using micrograms versus those actual doses, because you're trying to fit the model where it actually isn't quite getting the finalized projection.

Dr. Verstraeten: I think you have a point. I think one other point would be to just do it by 12.5 micrograms. I have done that and it is almost identical.

Dr. Rapin: At what age were these behavioral diagnoses made because that's a major issue?

Dr. Verstraeten: Most of them start from about two years of life and depending on the specific outcome, I think I have given you in the text you have received, the mean age for any of these outcomes. You will see that it varies. I think the speed, they are a bit younger. The attention deficit disorder, they are a bit older. But one thing is for sure, there is certainly under-ascertainment of all of these

because some of the children are just not old enough to be diagnosed. So the crude incidence rates are probably much lower than what you would expect because the cohort is still very young.

Dr. Walker: Following up on that, since you have a substantial part of the cohort which hasn't lived through the periods during which these diagnoses might be made most commonly, an elevate in association here could also simply represent a bringing forward in time of a diagnosis associated with a particular vaccination pattern. So something which would have been censored now moves into your observation period.

Dr. Verstraeten: That's absolutely true. I cannot differentiate between whether it's an overall increase or whether it's just bringing it forward. I agree.

Dr. Rapin: How did they make these diagnoses? You tell me that they're coded in the database, but how were the diagnoses made?

Dr. Verstraeten: What I am presenting to you now is just the results of the automated data. That means I don't know anything about how these diagnoses were made at this point. What we will present to you this afternoon is some of the results of the chart abstractions. I think at that stage we are in a better position, at least for some of the outcomes, to tell you how they were diagnosed.

Dr. Davis: Just to follow up on that, even when we get to that point what we are left with is sort of the real worldwide distribution of diagnostic patterns. So nowhere today or tomorrow will you ever hear that an analysis restricted to children that were carefully examined in the neuropsychiatric clinic. These are kids who are seen by regular old pediatricians who might eventually get referred

to a speech pathologist or attention deficit specialists, but the original coding is a pediatrician or a family physician who is making the diagnosis.

Dr. Verstraeten: Just for the sake of the presentation, could I go on? Because I see that we are going into questions about other issues.

This graph shows you a similar result for attention deficit disorder. One difference from the previous graph is that here the reference category is older children that received less than 37.5 micrograms of ethylmercury at three months of age. I did this because the numbers become so small that the estimates almost explode for some of these calculations. So for some of the disorders where the numbers are small, I have collapsed these three bottom categories and used that as a reference category. For attention deficit disorder we also have a suggestion of a trend. The test for trends is borderline, not statistically significant above one.

Dr. Chen: Go back one slide.

Dr. Verstraeten: I'm sorry, we skipped one. This is the result for autism, in which we don't see much of a trend except for a slight, but not significant, increase for the highest exposure. The overall test for trend is statistically not significant.

Now for the speed delays, which is the largest single disorder in this category of neurologic developmental delays. The results are a suggestion of a trend with a small dip. The overall test for trend is highly statistically significant above one.

Dr. Stehr-Green: I just want to point out that none of the point estimates for any dose level were statistically significant when you test

for trend. To what extent is that an anomaly based on the huge fact finding?

Dr. Verstraeten:

I think that is an important point that we will have to consider later on.

Here we do have one, but that's quite rare. What this represents is the overall category of developmental delays, of which I have excluded the speed delays because the impression we had was that some of the calculations were driven by this speech group, which was making up about half of this category. After excluding this speech group, this trend is also apparent in this group and the test for trend is also significant for this category excluding speech.

This is an example where there is rather a suggestion of a negative trend, however the test for that trend is not significant. There is a decrease for the highest category for cerebral palsy.

For the renal disorders, there is also not much of a trend, except for a slight decrease here for the highest category. The overall test for trend is non-significant, below one.

This shows you the results for premature children for the entire category of neurologic developmental disorders. What we see here is there is a very significant drop from children that were not vaccinated to children that received the minimum amount of Thimerosal-containing vaccine. After that there isn't much of a trend. The overall test for trend, which I think is in the text, is significantly negative. That is driven by this finding here. What happens here is that these premature children which are at high risk of having a disorder, or that is what we assume, are simply not being vaccinated and that results in an artificially high estimate for this zero group. However, what is also important to note is that after that we don't have much of a

trend happening there. That is one of the consistencies that we will have to discuss later on.

Now some results of when we tried to assess exposure by birth rates. We have birth rates for about 10% of the children in this cohort. That was done by linking the VSD to the states Birth Certificate files. It is only available for one HMO, for Group Health Cooperative, and it is only for about two-thirds of the children of that HMO.

What we have are the crude numbers again. Now I have divided the cumulative mercury level at three months of age by birth rates. Then I have categorized that exposure into different categories. I have tried to approximate quantals as much as possible, while keeping comparable categories. So it goes from zero to 14 because there are very few children with zero. Then 15 to 17, 18 to 20, 21, 23 and then above 23. The numbers in each category are comparable. These are the number of cases for the one category and two of the major outcomes. Then the rates. These are not adjusted for person time, it's just crude rates. It's just this number divided by this number which gives you this percentage.

What I have done for these two categories in the category of outcomes is first of all, I have looked at what is the influence of birth rates on the outcome itself? What we see is that for attention deficit disorder, this is not significant. Below one means the lower the birth rates, the more likely to get the outcome which is what we would expect for most of these disorders.

For speech, that does not happen. This is a strange finding. That the heavier babies in this cohort are more likely to have the outcome, and that is statistically significant.

For the overall category of developmental disorders, the estimate is below one, but it's not significant.

The next estimate I would like to point out is this one here. What happens if we divide the cumulative exposure by the birth rates? For attention deficit disorder, this estimate is near or a little bit higher than the one we had for the cumulative birth rate plain or not dividing by the birth rate. So it doesn't affect it very much and the confidence intervals overlap one.

What happens for speech, however, where this estimate for cumulative mercury exposure was significantly above one, it now goes below one. Although it's not significant, the significance disappears and the direction of the relationship becomes negative.

For the overall category of developmental disorders, we have a similar finding to the attention deficit disorder where this estimate slightly increases and the significance slightly increases also. However, we have to be careful in comparing this estimate to the one where we haven't divided by birth rate because we have a different scale. So it's not because it becomes somewhere around 7 or 8 to 25. That means an increase. More important would be the level of significance, which has only slightly increase.

Now a different approach. Instead of dividing the cumulative exposure by birth rate, is looking at the cumulative exposure and stratify the analysis on birth rates to see if that makes any difference on our findings. I have stratified by categories of 250 grams of birth rates. What happens if I do that the estimate, which I think before stratification was about 007 or 008, is hardly affected. Also for speech, after stratifying on birth rates, the estimate is not very much affected. So we have two quite different findings. If we stratify on birth rate it hardly affects the

estimate. If we divide by birth rates, it does affect the effect and I think we could have some interesting statistical or biostatistical discussion about this phenomenon.

For the overall category, stratification doesn't really affect the estimate very much, and dividing gives you a similar result if you take into account the different scales.

Now turning to the main limitations of this study. First of all, there is potential misclassification of exposure. Frank has mentioned that the hepatitis B birth dose can be missed for some children. I have looked at details of that and that is some of the additional analysis we could look at later on.

Thimerosal in haemophilus influenza vaccine originally were not shown, but we've been working together with people from the FDA and they have used the lot numbers that we have for each individual vaccine that is given. We have the lot numbers and we have sent those lot numbers to the people at the FDA and so far they have told us that less than 1% of the vaccines in our cohort, of the Hep vaccines in our cohort are Thimerosal containing. Less than 1% are Thimerosal containing, so everything all the others are Thimerosal free.

There is a difference in packages. If they are packaged in vials with 10 doses, they are Thimerosal containing. If they are packaged in vials with one single dose, then they are Thimerosal free.

Dr. Chen: Most of the vaccine that was used in the study contained Thimerosal.

Dr. Verstraeten: Right, so it was multiple dose vials. If it is single dose vials, it is Thimerosal free and hardly any of that was used in this cohort.

The birth rate information, we only have this on less than 10% of the cohort, so that information is limited.

There is the issue of using ICD9 codes for the outcome and someone already raised a concern about this.

There is the issue of medical care utilization factors. One of the main worries or one of the biases that we are particularly worried about is that the same parents that bring their children for vaccination would be the same parents that bring their children for assessment of potential developmental disorders. That could drive the estimates that we are seeing. There are a number of ways we have been trying to look at this, and we can look at that in the following discussions.

It's not just the parents, but it's also the health care providers. There is a potential that certain health care providers use more hepatitis B at birth and would also be more likely to diagnose some of the outcomes.

There is the issue that in the VSD we can only look at dose outcomes that come to medical attention. There is no routine screening of children, so it is only if the mothers bring their children for a problem that we will be able to pick it up.

Finally, for some conditions we didn't have sufficient power. That is particularly true for the renal disorders. We have very few cases in that category, so our bar is quite low.

There is inconsistency of our findings among premature infants. That is an important point.

There is the issue of excluding congenital and perinatal disorders. That has raised some concerns.

There is the question of variation and exposure. What does it mean exactly if a child has a low exposure or a high exposure? Basically because all vaccines have Thimerosal, it is a difference in being on time with your vaccination schedule. At three months of age some kids have received more vaccines than others, so what we are looking at is how well the children are following their prescribed regimen of vaccinations.

Finally, and this may be the toughest one of all, how do we know that it is a Thimerosal effect? Since all vaccines are Thimerosal containing, how do we know that it's not something else in the vaccines such as aluminum or the antigens?

In conclusion, the screening analysis suggests a possible association between certain neurologic developmental disorders. Namely Tics, attention deficit disorder, speech and language disorders and exposure to mercury from Thimerosal containing vaccines before the age of six months. No such association was found for renal disorders.

Dr. Rennels:

Do you have the data to show us of exposure at six months, or so fathered, by just saying three months because it is a nicer distribution?

Dr. Bernier:

Let me explain a little bit about how we structured this. We've presented this a couple of times and in the past people have raised questions. We have done analyses of these questions. We have presented the whole talk and the results of those analyses. We have found that it overwhelms everybody. So what we have decided to do is just do the regular abridged presentation, which he has just finished, and as you raise questions he will pull out of his question and answer bank. If you hit on a question that someone else has already raised, he will probably have the analysis. I trust that you will not think of all the questions

that others have raised, but on the other hand you will think of new ones that we haven't done already, so that's what they are getting ready to do. I think learning as you go may be a better way to digest this information.

Dr. Verstraeten:

What I have for six months is only looking at the exposure as a continuous variable. I have not reproduced this graphs. Unfortunately that is quite a bit of work, but if that is what you are asking I don't have that. I just have the continuous variables and these results were in the text you received. So basically what we are seeing is that for the ones that we are particularly concerned about, the speech and the overall category, these are also significant for those. It is also significant for the language delay.

For some, like attention deficit disorder, it only becomes significant at six months. For others like Tics, it loses the significance by six months. However, one thing you have to bear in mind, there is a high correlation between the exposure at three months and six months of age, which is what you would expect. Once the children get their vaccines early in the first three months of life, they are also more likely to get them earlier in the following three months.

I have a slide with these correlation coefficients, but it's probably not worthwhile in looking at those figures specifically. But what we see is that the correlation is very high between three and six months, but not as high between one and two or between two and three. So I would conclude that once the children are three months old, they are pretty much fixed in a high or a low category. Before that they can still change from a high to low category. It is not because they got their HepB in the first month that they will also get the other ones in the following months.

Dr. Rennels: But in fact by four months they may all be in the same category?

Dr. Verstraeten: They could be, but they are not. One of the main differences is the hepatitis B. Whether they got it or not, and those are already three doses. That's 37.5. Although you are right. There is pretty much two peaks and the difference between the two is the hepatitis B. Besides that, there is not such a wide distribution.

Dr. Gerber: I wanted to get back to the issue of medical care utilization as a possible confounder. You told us about a week ago that you were beginning to see some differences in the second year in one of the HMOs in terms of the number of office visits. Could you elaborate on that?

Dr. Verstraeten: Let me show you a couple of slides on what we have been trying to do. The first thing I looked at is the number of visits these children have in the first year of life during the exposure time. I have divided them into two different types of visits. Just a well child clinic which has specific ICD9 codes, or any other visit including those well child clinics. These are categories at three months of age. Then I have looked at the different categories of exposure to have an idea if there is a difference or not between the number of visits these children have and the different exposure levels. What this suggests is that as you go up the exposure levels, the number of well child visits increases, which is not really surprising because most of the vaccinations are given during those visits. However, although not perfectly, but there is also a suggestion of an increase, although it goes down here and then back up for the overall number of visits, so that is including a visit for any problem the child has.

However, when I adjust for these numbers, if I put this in the model as a co-variable or as stratified on it, it doesn't

change the estimates anyhow. It doesn't seem to make much difference.

Dr. Stehr-Green: Essentially it seems to me you may be calling the same variable, or the same characteristic two different variables.

Dr. Verstraeten: That's possible.

Dr. Stehr-Green: One thing I thought you might do is if these kids have siblings, you might take the average number of visits the sibling had and you could use that as a covariant. It can still be correlated with the visits that your study subject had, but it is not going to be calling the same characteristic two different things. There may also be other ways.

Dr. Verstraeten: Somebody has mentioned that before, what about siblings. We could look at that, but unfortunately I don't think we have the means in our automated data to find out who is the sibling of who, so that wouldn't be possible using the automated data, but that's definitely a great idea.

Dr. Walker: I'm troubled by this table. What you are telling us is the average child in these HMOs has 12 visits in his first year of life? Or 10 to 12. That number just seems a little large to me for an average number, and I am wondering what you are counting as a visit and that leads me also to ask you what you are counting as a diagnosis? I know these aren't claims databases, so it's not the diagnosis associated with every test.

Dr. Verstraeten: You are right, these are diagnoses, they are not visits. I'm sorry. These are visits. Unless they give them twice at the same visit, these are diagnoses.

Dr. Walker: So these are new diagnostic codes entered for a child, so a child could have multiple at one visit?

Dr. Stehr-Green: But those could include administration of vaccines, right?

Dr. Walker: Let me finish, please. So can you tell us the circumstances under which a code comes into the file? And you're counting it as an outcome. Maybe that's specific to each of the two HMOs.

Dr. Davis: First I would argue that this is probably normal. Even if they are visits, I would actually disagree with that this is above, because number one, you get your discharge from the hospital. You get your two week visit. You get your two, four and six month visit. Your nine month visit and your 12 month visit. Then you get your three colds in the first year of life. I think that's 11.

Dr. Walker: Well, that comes out to more than two by one month of life and you're averaging less than two.

Dr. Davis: I'm sorry, say that again.

Dr. Walker: The question is what do these codes correspond to? Are they a code given at the time of a visit with a health care practitioner or can these codes appear in any other context and still get into your file?

Dr. Davis: Yes, if they see an emergency room physician and I think for telephone calls we have some text strings. I don't think they get coded, so I think it's actually medical care utilization. They tend to be check box, so people would check boxes and then that gets coded in a different manner I might say, so that's how the diagnosis itself makes it into the automated file.

Dr. Walker: Now there are very few of these diagnoses which would actually result in only a single encounter and never again be the case of medical care. I would think that you could

get a lot of noise out of this system by looking at people who have had at least two visits with a particular code.

Dr. Verstraeten: You are talking about particular outcomes?

Dr. Walker: Yes.

Dr. Verstraeten: That is something we will talk about later. Yes, we have done that.

Dr. Chen: Why don't we show that slide now?

Dr. Verstraeten: Where they have been diagnosed more than once?

Dr. Chen: On repeat visits for the same...

Dr. Verstraeten: But I am not through with this medical care utilization, we will come back to that.

Now, for some of the outcomes, how many of these were diagnosed more than once? Autism, 40%, and there is a difference between the two HMOs. Speech delay, 37% and here it is higher at Group Health than at Northern California Kaiser. Attention deficit disorder, again the other way around, but they are pretty much in the same ballpark. The proportion of the cases in which the outcome has been diagnosed more than one. I think that was your question, correct?

Dr. Walker: Yes, have you done the analysis for each case?

Dr. Verstraeten: Yes. It comes back on the discussion also this afternoon of the chart abstraction. For attention deficit disorder, that is the same estimate except that confidence intervals become wider. The number goes down. For speech delay, actually the estimate slightly increases. This is at three months of age, so this compares to the 1008. The level of

significance, I'm not sure how that is affected, but basically it's pretty much the same thing. And I can tell you, although I don't have the figures in here, that for Tics there are too few, so I couldn't tell you. For language delay it's the same thing and for unspecified there are also too few because I think there are very few that come back twice. But basically for the ones where it was possible to do this, it was confirmed.

Dr. Walker:

Well, no, there is only two categories now in which you have enough data. That doesn't imply that the others are good.

Dr. Verstraeten:

I'm saying for the ones where it was possible to do this analysis, it would confirm what we saw. On top of that, I could go up twice, three, four, five and it would just increase the estimates basically, and that was only at Group Health.

Now going back to this medical care utilization, now I am looking at the number of visits. Just plain, the number of visits. Not just the first year of life, because another concern is that maybe the children of higher exposure come back more regularly later and have a higher chance of being captured later on. So what I have tried to do instead of giving you just these numbers, I have made just plain linear regression models of the exposure and the number of visits, to see if there was a linear correlation between the two.

What we see is that at Group Health, that appears to be the case. I have divided it by years of birth because I think it is important to keep in mind that there are these temporal trends. So for the different years of birth, at Group Health there is this trend. Really that the children with higher exposure are more likely to come back.

Now we get into the problem of mixing outcomes and confounding variables, because do they come back because they are sick or do they become sick because they come back many times? That becomes hard to differentiate.

At Northern California Kaiser, that trend is hardly there. It varies more around zero or it can even be negative for one year.

I don't think we should look at the significance of these numbers, but they just suggest that the trend is there.

This is the same, but just for well child visits and we see the same thing. That at Group Health there is a trend, that higher exposure groups have more well child visits. AT NCK that is not apparent.

These estimates are now using the number of well child visits in my proportional hazard models instead of the mercury exposure, and we see that for both ADD and speech delay, those two are significantly linked. So the more well child visits, the more likely to be diagnosed.

This is again looking at the mercury, but adjusting for the number of well child visits, and it doesn't affect the estimates. But again we have the problem we had before, that some of these variables now may be correlated and it's not obvious how that affects our estimates.

I hope this makes sense. Trying to adjust for these number of visits, but if this is very correlated to the exposure, that's not obvious if we can just do that. Anyway, we went ahead and did it and it doesn't really affect the estimates.

Dr. Oakes:

So that correlation will be taken into account in your confidence intervals?

Dr. Verstraeten: I'm not sure I want to say something about that.

Dr. Walker: Correlations which are you are accounting for when you control for confounding, so the fact of correlation is not by itself destructive of this.

Dr. Verstraeten: But at the same time if there is correlation, you may not be surprised that it doesn't affect your estimates.

Dr. Oakes: Well, it is true that if there is measurement error in either or both of these, which there almost certainly is, then it becomes less clear cut.

Dr. Modlin: If you could go back and have people take a look at slide 11 in your original presentation. My question...

Dr. Verstraeten: Before that, can I just finish up with the medical care utilization and then we'll get to that? Just to avoid jumping back and forth, if I can have the next slide on medical care utilization.

Something else we have tried to do, because we are thinking medical care utilization could be a link to socioeconomic status, and that could be another fact that is behind this, we have linked our data to 1990 census date, and then trying to assign race and income to the children. That is information that we don't have in our automated data, but we have been trying to do that by linking this. If we do that, we see this would be the racial distribution of our cohorts with the majority being white and then the second group would be Hispanic, followed by Asians and then blacks and a very few native Indians.

What I looked at here is what is the mean cumulative exposure at three months for these different racial groups, and they don't differ very much. The one that is different is the native Indians, but there was only three in this

category. So amongst the others, there is not much difference.

The outcome, however, can be linked. I saw that among the white group, they were more likely to have some of the outcomes of neurodevelopmental delays, which also is maybe not surprising. However, when I put these racial groups and stratify on it, it doesn't affect the estimates.

Also if I look at the estimates only within this group, it is also very close to the original estimate.

The next slide shows income. This would be household income and I've categorized them as follows. Between \$15,000, et cetera. Again, the mean cumulative mercury exposure does not differ between these groups, however, there is one group that is predominating the whole cohort. And again when I stratify on these groups, it doesn't affect the estimates, at least for the significant findings.

That is all for medical care utilization. We could return to your question.

Dr. Chen:

John, before we get to that. One way though to look at whether medical care utilization might be a potential confounder would be to look at other outcomes other than renal and neurological to see if we see the same kind of consistent trends. That might be useful before we jump to the other topic.

Dr. Verstraeten:

One other thing I did, what would happen if I just look at a few other outcomes that I don't think are related to Thimerosal. Am I going to see the same kind of trends? Maybe there is something in the data that I am not understanding. So I have selected a number of outcomes. First of all I have selected three outcomes among the most frequent outcomes, unspecified conjunctivitis, non-

infectious gastroenteritis and unspecified injury. These are the number of cases for these three outcomes. Unfortunately, I don't have the mean age. I haven't had the time to redo this and I hadn't written it down originally, because I think it would be important to better understand whether we can compare these outcomes. On top of that, we have selected two outcomes that we thought would be similar, also prone to the bias that the effect that the child has received that diagnosis somehow reflects parental concern. That not any child with these will be taken to a doctor. There is one code which is called worried well, which specifically states that the parent came with the child for a problem and the doctor said there was not problem.

The next one is flat feet, where we assumed there was a certain degree of parental concern needed to bring a child to the doctor for flat feet.

This is the graph for conjunctivitis. The same type of graph with the exposure categorized at three months. What happens is that here the zero group has a lower risk. It appears as if this group is just not being vaccinated and are not coming to the HMO. After that it is pretty much a straight line. Nothing much happens here once the child gets any vaccine.

They are all elevated compared to the reference category, but the trend is right here. Although that is significantly above one, that is a .1 risk. This is about only one-tenth of what we saw in the others.

Dr. Stehr-Green: Ironically you didn't show any of your Thimerosal related outcomes. Every exposure level above zero was skipped.

Dr. Chen: Because the ends are bigger.

Dr. Stehr-Green: Good point.

- Dr. Oakes: But the point you are making is that there is an artifact in the blood category there.
- Dr. Walker: That's one thing. This zero group has a problem. I think this zero group is a mixture of children. Either they are too sick to be vaccinated and that is a problem that happens with the prematures, or they just don't come to the HMO. They just get their vaccines somewhere else and then they are also not diagnosed.
- Dr. Davis: I might be very dense here, but they do get two polio vaccines.
- Dr. Verstraeten: Before the end of the first year. That's true. When we look at non-infectious diarrhea it is the same story. It goes up compared to the reference category, but after that there is hardly any trend and the test for trend gives the same result as the previous graph.
- Now in a way this one is also interesting. This is for injury where actually the trend now is down. There is a significant downward trend suggesting the more Thimerosal, the less likely to be injured. If one would try to explain this, is that the same parents who are concerned about having their children vaccinated are also concerned that their children don't get injured. That they are more careful.
- Dr. Stehr-Green: Again, point estimates suggest that none of those are significantly different from zero, so I think that's a rather spurious conclusion.
- Dr. Verstraeten: This graph is up and down. It doesn't suggest the same trend.
- Dr. White: Do we have vaccination rates for each one of these things? Aren't they greater than 90%? You showed to begin with

the vaccination rate of all the children in the age modes is greater than 90%, isn't it? You had included the vaccination rates of who gets vaccinated. Was it low or was it high?

Dr. Verstraeten: I don't know what the vaccination rate is.

Dr. Davis: Since that is tremendous over time. Maybe Ned, do you know the answer for NCK? At Group Health it was about 74% coverage by two years of age for a whole definition of what we are using today, but it went up to 91% very quickly. Like within two or three years after that. I don't think that is getting to what you were asking though.

Dr. White: I wanted to know if you were looking at these are parents who are giving their children these vaccines or not? I don't know.

Dr. Modlin: Maybe we could explore this question a little further because I think it is important. A couple of questions have come up. Actually Peggy's original question about exposure at six months of age also raises the same issue. That is if you look at your distribution of exposure that you showed in slide 11 on your original presentation, showing the frequency of exposure for numbers of each of the individual categories, you've got almost 2,500 kids that had no exposure. Zero exposure to mercury. About an equal number in the other two lower exposure groups. This is at three months of age at 12.5 and 25. Then of course your numbers go up by a factor of 10 or greater.

The comparisons here are critical because the zero exposure group is actually your comparison group, and since you're seeing trends in the data and it appears to be the trends that are bothering us the most, when you are comparing data in the higher exposure groups to the lower

exposure groups, these lower exposure groups at the relatively small numbers become very, very important.

My question is what is it about kids who get no Thimerosal, but still get two does of polio vaccine by a year of age that's different from kids who get exposure to the usual numbers of doses that we would expect if they are fully immunized by three months of age? My guess would be that these kids who are getting the lowest exposures are kids who are being immunized late. That's the only way in which they could get in the study. If they are getting their polio vaccines at the same time they are getting their DTP vaccine. So they are being immunized, and it may not be just the zero exposure kids. It may be those in the lower exposure groups as well who might fit into that category as well. So there is something different about them. That difference is probably very important.

Dr. Verstraeten: Let me show you some graphs. I looked at these kids at one year of age. How many of them were on time for their vaccinations or not, so let me show you.

Dr. Walker: While he is getting that out, the trend statistic isn't really being driven by that low dose group. You've have to look down to where the numbers are, so when he gives you a trend statistic it is really mostly averaged over the 37.5 to the 75, and the very dose ones are weighted more heavily because they are extreme, but still the numbers are so small. They are not much of the estimate.

Dr. Verstraeten: That is the other comment I wanted to make. When we have the trend estimate, we don't have a reference category.

Originally I had taken the high group as a reference category, but the first time I showed those results people were always trying to revert from below one to above one,

and that was so confusing, but then the graphs would go down. It was pretty much the same results, but then I decided to stick to these trend tests which I think are less bias because they are not fixed on one reference category.

But if we can look at this graph, what I looked at is that among these different categories at three months of age, how many kids end up being on time by one year of age? The end of the first year? So they would have their required number of DTP, HIB and polio, excluding hepatitis B here. What we see is that once they are at 37.5, almost all of them are vaccinated on time. The ones below these three categories, they are still about 50% and strangely enough this one was even lower. There are still about 50% that get their vaccines on time. There is another 50% that doesn't get them on time and this is the one probably to worry about.

This is the same thing including hepatitis B, and not surprisingly those figures are increased a bit where there are a higher number of children who do not get their hepatitis B by the end of the first year of life, or don't get their entire vaccination schedule by the end of their first year of life.

Dr. Oakes: I'm missing something here. If they are getting their vaccinations on time in the upper part of this, why is there any difference in the exposures at all?

Dr. Verstraeten: There are several possibilities. In this they get their vaccinations on time by the end of the first year of life, so they might have gotten it, not before three months, but after three months. Then there is also a difference between DTP-HIB combined or separate. That makes a difference of 25 micrograms. That is something Phil will talk more about this afternoon.

- Dr. Stehr-Green: This brings my level of concern even higher. It may not be an issue of confounding we are dealing with, it may be an issue of bias. Whereas these kids who aren't getting vaccinated in the first three months of life, they are just essentially dropping out. So not only are they not getting exposed to Thimerosal, they are not getting an opportunity to be diagnosed with any of these other outcomes. Yet they are still in the cohort because they make their entry criteria of having two polio vaccines, but they are not having enough visits to get either vaccinated and therefore exposed, or to be seen and get diagnosed. So it seems to me it may not be an issue of confounding, but we have to think about an issue of ascertainment bias.
- Dr. Verstraeten: It's possible, but we are also not sure. We don't know why these children don't have visits. Maybe they could come, but they don't come for some reason.
- Dr. Stehr-Green: But the question of why may be irrelevant. I'm saying that may be what is driving some of your observations.
- Dr. Modlin: They have visits, they are just delayed. They are getting visits because they are getting their two doses of polio later on and ultimately becoming fully immunized.
- Dr. Johnson: But how would that explain the alternative diagnoses? The trends we see there. That explanation would have to apply to both the mercury plausible outcomes as well as those alternative.
- Dr. Walker: You could address the criticism pretty easily and without much cost by simply truncating your lowest exposure levels since you don't have very many people anyway, and taking as your reference group a typical.
- Dr. Verstraeten: Right, that is something Phil will talk about this afternoon.

Dr. Mawle: I'm still on this same graph, with your different levels of exposure between 12.5 and 50, you have two different ways of getting there. You can either get two doses of hepatitis B, which presumably would occur at two different levels, or you can get them all in one. Now one of the problems you have with Thimerosal is you don't know what that does to the actual blood levels of the body blood, but presumably if you've got them spaced it would be different than if you got them all at once. Did you analyze them separately?

Dr. Verstraeten: I know, that raises the other issue. Exactly like what you're saying, the timing of when they get this may also be important and maybe this comparison is not perfect because some of them got it at one month or two months, and it's pretty hard. What I have tried to do is like stratify on what they got before that, but then you start mixing up things. It becomes quite confusing.

Another possibility is giving it different weights depending on when they get it and the later after birth, the less weight you give it, et cetera. There is different ways to go about it, but I think at a certain point it becomes a bit too complex or a bit too confusing, although you can still try to do that. But to sort of understand what's going on, it gets a bit too mixed up. In a way it's possible to do that, but the variation decreases a lot, too, if you start doing that and if you start stratifying about what happened before, about what happens afterwards, you loose.

Dr. Mawle: I was concerned. There's a big difference between getting them all at once and getting them spaced presumably at least a month apart, and when you look at the levels we're talking about, which is a chronic exposure versus the acute exposure, those three categories are not comparable. At least they're presumably not comparable. And I think that the NIH studies are supposed to be addressing some of

those issues, but I don't think that at this point we can truly say that all of your 37.5 for instance are equivalent.

Dr. Guess:

I'm not entirely convinced by the analyses showing no trend in these other diagnoses because for example, the gastroenteritis and the conjunctivitis would be things that you would think the parents would probably bring children in for, whereas some of these developmental things, particularly the more subtle ones, may not be. In other words, a profound developmental, yes, but a subtle one perhaps less so. So it would seem, to me anyway, that to rule out the issue of the ascertainment bias, one might need to examine other kinds of diagnoses not thought to be associated with Thimerosal and which may be things that parents may not bring people in for. So I think it's a good line of reasoning, but I'm not sure it's been entirely put to rest.

Dr. Verstraeten:

Could I have slide 32, I think that addresses that question. I'm sorry, 31. I mentioned these other two diagnoses, the flat feet and the worried well. I haven't showed you the results for those. What we have for those two are the estimates. For the worried well and the flat feet, both of them are non-significantly different from one. They are both below one. The finding is not significant. That's at one month and at three months of age.

The last category, maybe I will talk about this now because it's good to be aware of this. There is analysis we have done where we compare the children that got DTP-HIB in the combined vaccine or DTP-HIB separate, which is a difference in exposure of 25 micrograms. But basically we assume that these children are comparable. They get the same number of antigens. They get the same number of vaccines. They come pretty much at the same time for vaccinations, except that one gets the vaccine in one shot and gets 25 micrograms less than the other children. So

when we do that, at least for these outcomes, we see that it's all rather centered around one and none of this is significant.

If we do that for the other outcomes...

Dr. White: Would that make it confounded by different HMOs? You were talking about that back here where the HIB combined was used in one HMO.

Dr. Verstraeten: Right, this analysis is limited to one HMO, not the other one. Absolutely. Anyway, I think that was something in the text that you have received. Does DTP-HIB combine on separate analysis? The original test that was handed to the people.

Dr. Myers: The one they got in the mail?

Dr. Verstraeten: Yeah, the one they got in the mail. Basically what we found when we do that is that for most of the outcomes, or for all of the outcomes, none of the estimates are significant. Most of them are above one, but none of the findings are significant. However, the power of this analysis is limited because it's basically only in one calendar year that it happens. That vaccines were given. Some kids got the combination vaccine or some kids got the vaccines separate.

However, among prematures, that becomes significant and we get relative risks up to two and three, whereby the ones who got more Thimerosal are at higher risk than the ones who got the combination vaccine, so about 25 micrograms less than Thimerosal. However, the number of children in this analysis is quite small and that result is quite sensitive to small numbers.

Dr. Snyder:

This issue of ascertainment biases is obviously something of great concern. With regard to that there is a piece of data I haven't seen yet that you may have looked at which has to do with looking at the proportion of children by level of mercury exposure who remained in the HMO at 18 months or two years or three years. At points in time at which the cases would be ascertained. If there is no difference in the proportion, it gives you a higher level of confidence that there is something there, whereas if they don't remain in the HMO it just exacerbates the concern about ascertainment. Is that analysis condoned?

Mr. Verstraeten:

That has no meaning.

Dr. Davis:

The analysis as it is set up takes that into account though because people are censored at the point that they drop out of the analysis, so basically at any given age in their life, let's say at two years of age, you are comparing people who are put on the analysis based on their exposure category, then they are followed up for the outcome and then censored when they drop out. So we are not really concerned about people who disenroll from the HMO.

I think maybe what I should do is just suggest to restate to your concern. They are sort of dropping out from health care seeking behavior, but remaining in the HMO. Maybe that's what you are getting at.

Dr. Rhodes:

The point that Dr. Davis makes about censoring is fine except people who are at higher or lower risk are more likely to be censored, then it's still a problem.

It does turn out that the kids who are in the low group at GHC actually leave the HMO faster than kids who were in the higher groups, at least in terms of disenrollment dates. Now as to what they're doing in terms of their medical care before that, that's not in question.

Dr. Stehr-Green: With regard to ascertainment bias from a general pediatrician's point of view, the outcomes that have been produced by this study, a neurodevelopmental and a neurobehavioral outcome in children no older than five or six years, can be very dependent upon the concern of the parents. Particularly speed delay. There are parents that will tolerate tremendous variations in speed and language in the first three to four years and pediatricians rarely see children or evaluate children speaking in their office to the extent that they can make that diagnosis. So I think you have a real bias in the interest of a parent to make this diagnosis, and how you can use that in comparison to Thimerosal levels, I don't know. But I think it impacts on your conclusion tremendously.

Dr. Verstraeten: I agree. That's the main bias we have a problem with. The only remark I would like to make is that we always assumed that concerned parents would also have their children more vaccinated. I am not sure if that is something you can just assume, but that's the underlying assumption that we are making.

Dr. Johnson: There are a lot of questions remaining and I think we'll have to decide during the lunch period how to deal with those. If we do not break now, we risk not having any lunch at all, so we have to start with that. We'll be back at two o'clock.

Thank you Tom and also Bob Chen so deftly managing the slides.

Dr. Verstraeten: I did manage to find the slide I couldn't find before lunch, we'll start the afternoon session with that one.

Dr. Stehr-Green: If you can take your seats in the back please, as we are limited for time.

Dr. Verstraeten: These are the risk estimates by comparing the DTP-HIB separate to combined, which is a difference of 25 micrograms of ethylmercury and the combined with the lower mercury content is the reference group and these are the findings.

As I mentioned before, almost all of them except for this one are above one, however, none of them is statistically significant.

Dr. Rodewald: What kind of end are you talking about? How much...

Dr. Verstraeten: It would be about between one-third and a half. I apologize, I didn't put that, but it's anywhere between one-third and a half off the total sample size. So say for speech that would be about 500, more or less.

Dr. Rhodes: Tom, if you look at it, is this limited to Northern California?

Dr. Verstraeten: Absolutely, because in Group Health they didn't use the combined vaccine, so it's only Northern California.

Dr. Rhodes: The ends here can be very confusing because of the way the models have been fit in a very stratified fashion based on month of birth.

The switch over from a separate DTP-HIB to a combined HIB at NCK was done very quickly over the course of a couple of months, or at least as it appears in the data. So that when you stratify by month of birth, you essentially throw away all those kids that occur before the switch and all those kids that occur after the switch has been completed, although as has been said, there are some possible miscodings in kids who appear to have the wrong version after the switch. So that there are sort of two problems here. One is that you may start with say 1,000

cases, but if only 100 of those occurred during the switch, you're working with 100 cases and not 1,000, and if cases occur after the switch, but they are miscoded, then they will inappropriately enter into the analysis.

Dr. Sinks: Just interpret the relative risk again for me. Is that again using the model of one microgram, a unit change of one microgram or what?

Dr. Verstraeten: No, this is the relative risk as you are used to seeing it. If you are in one group compared to any other group, then your risk is say for 313, it's 1.5 and it's not by micrograms of mercury, no.

Yes, the difference is 25 micrograms, but it's not divided by the micro...yes please.

Dr. Guess: I just wanted to clarify the question that Phil raised. Am I to infer then that the sample size is restricted to the children who were getting both vaccines during the time period when both were being used, and it doesn't include the people when it was all combined or all non-combined? Am I understanding that correctly?

Dr. Verstraeten: Well, actually it was including all of them, but the way the model works, the way that it's stratified on month of birth, there isn't much that can be compared in any of the other months, so those weren't, wasn't very much to them.

Dr. Guess: Okay, I understand. Yes, thanks.

Dr. Brent: I would like to go back to the design with regard to the pharmacokinetics. The fact is that in the introduction it really is unclear as to whether this is a water soluble form or whether it's organic. In other words they say that. I don't know whether it's on to the other, but the point is that if you administer these doses over the period of an interval

of weeks, we don't have any evidence that the level is rising. In other words, if the exposure is changing. If the half life is like two weeks or 14 days or 18 days, by the time they get the next injection you're back to the background level. So that the whole idea that you have successive increases in exposure may not be true. The fact is that six months the blood level of mercury may be exactly the same as it was after the first dose. So that all those calculations of adding up the doses, if you have 62.5 micrograms of mercury, may not be true.

The second thing is, when you talk about neuro-behavioral effects, you're talking about what we all deterministic effects. They are thresholds. I don't know about ethylmercury, but methylmercury, the threshold for neuro-behavioral effects in like the rhesus monkey and in many animal species is way above the exposures that these infants are receiving. It's in milligrams per kilogram, not micrograms per kilogram. So all these levels, whether there is a dose response curve or not, may be below the threshold for producing any neuro-behavioral effects. So I think it would be very important to get the pharmacokinetics out of the way to find out what are the blood levels or the tissue levels of the ethylmercury in the infants over this six months period.

You know, all these calculations, statistics and re-analysis, if it's not based on firm pharmacokinetic exposures is not very easy to interpret.

Dr. Verstraeten:

Thank you. I think to answer the first part, what we have been saying is that this is the cumulative amount that these children have received. That does mean that at three months that would be related to their blood levels or hair levels or whatever. That this accumulates in the blood. That's not what we've been saying.

All we have been saying is that this is the amount they received. We know that's true. If that amount is accumulated in the tissue or in the blood, that we are not aware of. And as you are saying, as long as we don't have the pharmacokinetics at mercury, there is no way we can assume one way or the other. So we can only work with what we have, which is the amount that they have received.

Dr. Brent: Because the most important thing with the biological effect is the dose that the central nervous system or the developing cells are receiving. If you never raise the dose and the dose is always below the threshold, then you don't have a biological effect even possible.

Dr. Verstraeten: Right, but at this point there is nothing we can say about the actual dose.

Dr. Weil: Bob, you are assuming a threshold. The hypothesis here sounds like it's an exposure dependent related, dose related, and you don't know what is below the threshold you are referring to, which was an animal derived one.

Dr. Brent: There are two kinds of effects from a so-called toxicological viewpoint. One is called a stochastic effect, where the dose goes to zero. In other words there is no dose that presents no risk. And the second is the so-called toxicological S-shaped curve where the dose is S-shaped, and when you get down to a certain level the effect is no different than the controls.

The only diseases that have a stochastic effect where the dose goes to zero are those diseases that can produce by changing a single cell, and those are malignancies and genetic disease. Those are the only two diseases that have a risk from let's say a mutagenic exposure that goes to zero. Everything else has a threshold because it is a multi-cellular phenomenon. You cannot produce learning

disability by changing one neuron developing, or autism by changing one cell in the cerebral cortex. It's a multi-cellular phenomenon. Therefore it has to have a threshold.

I don't know what that threshold is, but based on the methylmercury data it is far above any dose that we are presenting to infants in these studies.

Dr. Verstraeten:

Two issues. First of all, like you say the threshold is established from methylmercury. I think we should avoid a discussion of how do we compare methylmercury to ethylmercury. I think that would take us very far.

Secondly, we are talking about biologic plausibility, and I would ask that we reserve that for later on when we have the appropriate time to discuss those issues.

Dr. Weil:

I think what you are saying is in terms of chronic exposure. I think the other alternative scenario is that this is repeated acute exposures, and like many repeated acute exposures, if you consider a dose of 25 micrograms on one day, then you are above threshold. At least we think you are, and then you do that over and over to a series of neurons where the toxic effect may be the same set of neurons or the same set of neurologic processes, it is conceivable that the more mercury you get, the more effect you are going to get.

Dr. Brent:

For every dose you give, it's gonna get above the threshold, because what it is, below the threshold the recuperative powers of the tissue enables not to respond in a negative way. You have to be careful if you keep forgetting about the importance of dose. I don't care whether you give it one time or four injections over a period of six months. It's whether the level below the dose that affects the development of the central nervous system, they're not going to have an effect.

Dr. Verstraeten: Excuse me, I understand all this, we've only got 15 more minutes to have the discussion.

Dr. Johnson: Yes, let's hold this. Just put it aside for a while and we'll come back to it.

Dr. Verstraeten: There is just a number of slides I would like to show because I think they have their own importance.

Next slide, please. I mentioned at a point that it's very hard for us to differentiate whether it is Thimerosal effect or anything else.

What I have done here, I am put into the model instead of mercury, a number of antigens that the children received, and what do we get? Not surprisingly, we get very similar estimates as what we got for Thimerosal because every vaccine put in the equation has Thimerosal. So for speech and the other ones maybe it's not so significant, but for the overall group it is also significant. So that is very difficult to distinguish.

Here we have the same thing, but instead of number of antigens, number of shots. Just the number of vaccinations given to a child, which is also for nearly all of them significantly related.

Dr. Myers: Tom, just on the number of antigens, did you add in the other antigens that were dropped at the beginning?

Dr. Verstraeten: Yes, I added polio which was basically the one that was missing. It doesn't change, no.

Dr. Guess: What are the units here?

Dr. Verstraeten: The number of antigens.

Dr. Guess: So this essentially in a 7% risk per antigen, and an antigen is like in DPT you've got three antigens.

Dr. Verstraeten: Correct.

Dr. Egan: Could you do this calculation for aluminum?

Dr. Verstraeten: I did it for aluminum. Actually that was the last thing I did last night before I left the office. I just did it for NCK because for Group Health it would have been more difficult to program. Actually the results were almost identical to ethylmercury because the amount of aluminum goes along almost exactly with the mercury one. There is one vaccine, HibTITER, that doesn't have aluminum, but then if they get a HibTITER, they get a DTP and the DTP has aluminum. So they are almost identical.

Dr. Egan: You were doing these as the number of antigens, not as the number of shots? Because the more shots, the more Thimerosal.

Dr. Verstraeten: Yes, I did both, number of antigens, number of shots. The first slide was the number of antigens, the second was the number of shots.

Dr. Egan: So in other words, some of the children are missing their vaccines then? Or at least for that time period.

Dr. Verstraeten: Yes, which is the same as before.

Dr. Sinks: Absolutely there is a lot of correlation or co-linearity between this analysis and your primary analysis with mercury, but in terms of evaluating the confounding, it would be nice to see what happens with the risk estimate in the model that's showing these things, so you can actually see is it blowing up on you. What is actually happening. How co-linear are they.

It is not surprising at all that we are seeing this. The size of the relative risk is obviously different because you are looking at different units and you can't compare one microgram of mercury versus one antigen. But it would be nice to see in a model both of these values, the relative risk at the same time.

My guess is that what is happening, I wouldn't expect both of them to remain statistically significant.

Mr. Verstraeten: You mean if I put both at the same time? I didn't try that.

Dr. Sinks: Oh, okay, you didn't have them in the same model.

Dr. Verstraeten: No, these are separate. Absolutely. No, I just showed this to illustrate that with this data it is pretty impossible to differentiate.

The only option we have is the DTP-HIB combined or separate. That is the only one where the aluminum is identical, the number of antigens is identical. Only mercury is different at that point.

Then the last slide I wanted to show, there was a question of if there was any way from this data that we could estimate what would happen in the future if there is Thimerosal-free HepB and Thimerosal-free haemophilus influenza vaccine and only DTP has Thimerosal.

What I tried to do is I took out of the cohort those children that increased their Thimerosal amounts by 25 micrograms between one and three months of age, which is when they have already received the HepB and when they have received their DTP and Hib. Those are the estimates right here. So those are the children that between one and three months of age, they have increased their mercury amount

by 25 micrograms, which is what would happen if DTP would be the only vaccine with Thimerosal.

None of these estimates are significant, however, the sample size has gone down quite a bit. I'm sorry, I don't have the numbers here but they are around 100 to 200. They are not very high.

The second column would be the same scenario, but now at six months. Assuming they have received two additional DTPs, so between three and six months of age they have increased their ethylmercury amounts by 50 micrograms. If I do in this current cohort with all its limitations, because there is also the HepB that exists in this cohort, I can't really take it out. It is significant for this one disorder which is language delay and it is quite high. Together with that, speech or language delay which is a combination of these two disorders, also becomes significant.

The overall group is borderline, not significant. Basically what one could say, if you can assume that this is a valid analysis, it doesn't give you complete security. I mean there is still a problem at this level.

Dr. Staub:

I am a little confused. In this analysis these children would not have received a hepatitis-B at birth dose, is that correct?

Dr. Verstraeten:

They can have received. I have done it irrespectively of whether or not they have received that.

Dr. Staub:

So I guess I will ask my question that I talked to you about before lunch, which was in the pre-reads that you sent us there was a table which had your statistically significant results in it, and language and speech delay were significant at one month of age and essentially carried that significance through the rest of the analyses.

And then your graph that shows the relative risk increasing in speech delay actually has a dip at 25 micrograms. When I saw your slide 11, vaccines contributing to mercury distribution at three months, the scenario for 25 micrograms, actually 75% of that group does not get a Hepatitis-B dose at birth, and 25% does. I guess I just wanted to make that comment that it appears to me as though more work needs to be done on the Hepatitis-B dose at birth scenario.

Dr. Verstraeten:

I don't know if people managed to follow that because we discussed this before lunch.

What happens in the graph for speech disorder is that you have sort of a dip in the third category of 25 micrograms, which is something we were rather puzzled about. One possible explanation would be that in this 25 microgram, the majority of those children received the DTP-HIB combined and received no Hepatitis-B, so they were a little bit at a lower risk because they didn't received that Hepatitis-B in the first three months, also in the first month, and that would be a possible explanation.

However, some of the analysis at three months, I have done them stratified on whether or not these children received hepatitis-B in the first month. For some of the outcomes, this relationship still persists. Meaning that you cannot explain it entirely by the hepatitis-B effect in the first month. Also what Phil will say, it's not dependent on this hepatitis-B, so you can't entirely blame the whole thing on the hepatitis-B in the first month.

Dr. Johnston:

I wanted to go back to power point 18, it's page 9 on the hand out we got this morning. Why is there such a difference between the Group Health Coop and the Northern California Kaiser? Even at no injection or no cumulative mercury exposure, there in speech delay there

is almost a doubling almost all the way through. What's the explanation? It is listed as 18.

- Dr. Verstraeten: For the rates.
- Dr. Johnston: The difference in rates.
- Dr. Verstraeten: I don't know, I'm not sure. Why the incidence rate for speech delay is so much higher at Group Health as compared to Northern California Kaiser.
- Dr. Johnston: If they are doubled all the way through the extent of the exposures.
- Dr. Verstraeten: One thing that Bob just mentioned is that at Group Health they have their own referral center for speech and learning disabilities, and it seems that sort of facilitates the General Practitioner or the Pediatricians to more easily refer the children because it is within the HMO and it is probably taken care of. So that might be one reason why more of these kids are picked up. That is one hypothesis. I don't know if the people at NCK want to say something. Ned?
- Dr. Lewis: Ned Lewis, NCK. That's right, and also the speech therapy is not covered.
- Dr. Verstraeten: So that is one hypothesis. It appears not be covered at Northern California Kaiser.
- Dr. Johnston: But it makes you a little worried about the endpoint? The outcome? When it is so different from location to location.
- Dr. Verstraeten: Right. Also what we are doing more and more is the analysis separate for each HMO because we sort of realized that we can't compare those two.

- Dr. Oakes: I am wondering if it is feasible to stratify the analyses by pediatrician? By diagnosing pediatrician?
- Dr. Verstraeten: We haven't been able to do that. One thing that Phil is also going to mention is stratification by health facility at least, and that we can do only for NCK because we know at each facility a diagnosis was made, and that also plays a role. Then of course there is still the level of the pediatrician which we haven't been able to reach.
- Dr. Oakes: Is that a feasibility problem, a data problem or a conceptual problem?
- Dr. Verstraeten: I'm not sure, Frank, if you have an idea on that or want to comment on that. Any or all feasible or if you have an idea about that?
- Dr. Rhodes: I think at NCK I was able to assign a sort of usual clinic to most of the kids. I think going beyond that is really impossible at the level of data we have now. I'm not even sure whether in these clinics if there is the pattern of the same pediatrician seeing the same kid over and over or whether it is just who is available, and I'm not sure how that goes at Group Health or NCK.
- Dr. Rapin: I wanted to know if the endpoint, this diagnosis of language disorder, autism, Tics or whatever, was it done just once? I mean, to enter your statistic, if the kid had that diagnosis once at whatever age he's in for that diagnosis? How does it work? I don't understand.
- Dr. Verstraeten: The main bulk of the results I have shown you is just once. However, we have done it specifically for the ones that were diagnosed more than once. That is one table I have shown you, but only for a few outcomes have we done that. And also it only works for those that have quite high

numbers because it is only like less than half that comes back or are diagnosed twice. But in general it is only once.

Dr. Davis:

I want to actually start off my talk in a little unusual fashion and acknowledge the amazing amount of work Tom Verstraeten has done. I am not sure if people realize that this has been sort of a full time occupation. This is really a remarkable piece of investigative and analytic work that Tom has done, with help from others certainly.

That said, I am going to talk today and try to address at least one of the concerns people have. Which is that so far all of the analyses done to date have been based on the automated codes and yes, we have used different slices of the automated coded, but we are still using the automated codes.

So in fact we, over the last three weeks, have done a chart review of over 1,000 charts at Northern California Kaiser and at Group Health, specifically looking at children with speech delay, autism and attention deficit disorder, to try and answer this particular question. Which is how good are the automated codes? And then specifically are they all similarly accurate? That is, is speech delay automated code as good as autism? Is it as good as attention deficit disorder?

Then further on, does accuracy differ between institutions? How can we use this information in terms of children being referred to specialists and speech therapy? And what kind of role can we assign to the history of past and present otitis media, and the role of other conditions in how well is the use of automated codes.

Then at the very end I am going to show you some brief data where we have actually, or Tom, has actually redone the entire analyses that you have just seen using only cases

that were verified as being “real cases” and using different definitions of real.

There is at least an hour’s presentation here and I know I have 15 minutes, so I am going to go through this rather rapidly.

When we look at speech delay in particular, we find that, believe it or not, some times it is not even mentioned in the chart and this is just a recurrent theme. It is not coming as new to those of you who have done chart review. Of the 577 cases of speech delay, we found it mentioned in the chart 560 times, or 97%. Of the entire group, 91% were referred to a specialist, so 91% of everybody who had an automated diagnosis of speech delay actually was referred to a speech specialist, and of the original group 75% were confirmed as having a speech delay by a speech specialist. Then a smaller percentage were referred for speech therapy.

There is a question we will see later on, but for this diagnosis at least, speech delay being mentioned in the chart does not vary between HMOs.

In terms of the proportion that gets referred to the specialist as we saw previously, at Group Health there seems to a slightly increased rate of children who are referred to a specialist, and again even though this is a small difference it is probably related to the fact of the easy availability of a language pathology center that is specially designed to take care of these children.

Of those that are confirmed by a specialist, this is the original number we started out with, but a higher proportion are confirmed by a specialist at Group Health than at NCK. And as Tom, Frank and I found at Group Health, when you were referred to a specialist it was almost

a done deal that you were almost in fact confirmed by this specialist.

You will see that there was a much higher proportion of children at Group Health who are referred for speech therapy than at Northern California Kaiser, and these of course relate primarily to coverage issue.

Now in terms of a search for pre-disposing factors, this is actually going to be important in what I will talk about tomorrow, but I will mention it today and put a little seed in your mind. Which is that serous and chronic otitis media, by history being mentioned by the pediatrician or the specialist, was present 38% of the time. It was slightly more present among Northern California Kaiser patients than at Group Health Cooperative.

Serous otitis media or chronic otitis media being actually present at the time of the first visit was present less than 5% of the time among these children, and only 4% of the children actually had a hearing loss that was tested and confirmed, either at the present time or in the past time.

This again speaks to an issue I will raise tomorrow, but it was interesting to us how often other possible pre-disposing factors for speech delay were present and recognizable on the chart. Bilingual language in the household, mental retardation, attention deficit disorder, developmental delay or other developmental disabilities, overall approximately one out of four children who had speech delay had one of these pre-disposing factors. And of course simply the presence of one of these pre-disposing factors should not lead us to attribute the speech delay to the pre-disposing factor. It actually is all tied up with the relationship between the pre-disposing factor and the speech delay itself.

In terms of autism, there was a code and the code occurred 120 times and autism was mentioned in the chart 92% of the time. It was actually coded. Of these 110 that were mentioned, 105 in fact were referred to a specialist. I have a feeling the reason that they were not all referred simply refers to the fact that some people were probably censored from the data set before they could be referred or they disenrolled and enrolled in a different health care plan. Of these 105 that were referred to a specialist, 99 were confirmed by a specialist and 6 had some other diagnosis. I imagine that would be suspicious for autism, but in fact turned out to be something else.

There were really fairly limited differences between the two sites in terms of the predictive value of the autism diagnosis. When it was mentioned in the chart, around 90% of the time it was found in the chart at both sites.

At Group Health Cooperative, when we saw a patient who had autism mentioned, 92% of the time they were in fact referred and very similarly at Northern California Kaiser. Note the very small number here, so it would be one more or less case would actually affect this percentage point by eight percentage points, so I consider these equivalent.

In terms of confirmation by a specialist, again 92% of the patients at Group Health and 81% of the patients at Northern California Kaiser had the diagnosis of autism confirmed by a specialist.

Now I think we get into somewhat different findings, which are attention deficit disorder and attention deficit hyperactivity disorder. I don't think the findings here, the fact that they diverge from the previous two diagnoses, is in fact going to surprise anybody. ADD was coded 348 times, and in fact we only found it 249 times, 72% of the time, which was somewhat less than we had previously. It

was referred to a specialist quite a bit less, 49% of the time, and was confirmed by a specialist even less, 31% of the time. So the predictive value of these codes is only 31%.

Our ability to find ADD if it was coded was similar between sites. But in fact, being referred to a specialist really diverged. Marty, you probably know about this. Who is the specialist at Group Health? We have somebody who has in essence devoted his entire life to the treatment of ADD and I thought he worked with you on the practice parameter for ADD.

Dr. Myers:

They do have a center for it.

Dr. Davis:

And that's what I'm getting at here. They actually have a center for the diagnosis and treatment of ADD and ADHD and I must say being a pediatrician 10% of my time, it would be a joy to have a center where you can easily send children for the proper diagnosis and care and this is not available at Northern California Kaiser, and probably accounts for the difference in predictive value of this particular diagnosis.

The diagnosis is confirmed more frequently at Group Health, probably using some standardized criteria.

Just to wrap up this section on the confirmation of automated diagnosis, how good are the automated codes? I would say for autism, the predictive value of an automated code is 81% and I rate that as very good, using my completely subjective rating code that I came up with last night. It's good for speech delay, with a predictive value of 75%. And it is also poor to fair, that is if ADD is in fact coded, you only have a 31% chance of finding a confirmed diagnosis of ADD or ADHD in the medical record.

Does the accuracy of these codes differ between the institutions? I must say that I did not find any consistent differences, although one can make an argument that the accuracy may differ for ADD, ADHD and probably relates to center differences or the availability of specific centers and perhaps reimbursement practices.

I think I am just simply going to specify that. To my take, the speech delay attributed to hearing loss or otitis media problems, by our chart review we found on 4.2% of children whose speech delay was directly attributed by some medical examiner to hearing loss or otitis media problems in the past. I would have to say that the medical record review is of tenuous value for this purpose and simply not worth it to go after this particular historical facet.

Now you are probably all wondering we did this medical record review, how are we going to use the results? Well, in fact we have replicated the analyses. Let me walk you through it because there is a lot of data packed on two or three slides.

This is the relative risk for speech delay per microgram of exposure. So we are back to that unit or metric of exposure. This is all cases with the rejoinder that Dr. Rapin mentioned. This is now the relative risk for all cases of speech delay, where the cases had to be seen at least twice. So it is not the ones that came in, that were evaluated and were felt not to be speech delay.

Per microgram of exposure by one month of age, the relative risk was 1.018 with the confidence interval as shown here.

Now one might imagine that would just disappear once we actually confirmed these diagnoses from chart review, but

in fact it did not. You see if the diagnosis was mentioned in the chart, the relative risk increases ever so slightly. I'm not going to get into an argument of whether that is a true increase or not. As a matter of fact it did not disappear.

In terms of when we cut it a little finer and insisted that all patients had to be referred to a specialist or had to be confirmed to a specialist, in fact the relative risk was down 1.026 with confidence intervals of slightly tighter than seen originally. Which is actually kind of interesting because the power fell somewhat. The power fell actually about by 34% here, so the fact that the confidence intervals tightens up a little bit in the face of a fallen power is a little interesting.

When we look at exposure by three months of age, again using the prior definition of all cases, relative risk of 1.013 and if we limit it to children whose diagnosis is mentioned in the chart, children who are referred to a specialist or children who were confirmed by a specialist, the relative risk stays about the same, with a relative risk of 1.016 among children who we were measuring the exposure at three months of age and whose diagnosis were confirmed by a specialist, with a confidence interval of 1.006, 1.026.

Now this other information that we collected. Again, we are just comparing it to the standard here.

If we are looking at the exposure at one month of age, the diagnosis of speech delay was in fact mentioned in the chart. We've excluded those children where the speech delay was attributed to a past history of chronic serous otitis or chronic otitis media, and we have excluded all those children who had mental retardation, bilingual family, attention deficit disorder and other contributing conditions. The relative risk in fact increases 1.025 with a confidence interval as shown.

If we limit it to children where the diagnosis was mentioned in the chart and we excluded any children with past otitis media, where the hearing loss was not attributed to the past otitis media. It's just children with a history of past otitis media, the relative risk is similar to what was seen just previously.

Now we are getting finer. If we eliminate the children confirmed by specialist, excluding those whose speech delay is attributed to past chronic otitis media and we are also excluding children who have other contributing conditions. The relative risk is now 1.031, confidence interval as shown. And if we are limiting it to the even smaller group of children that are confirmed by a specialist, and excluding any children with a past history of frequent otitis media, the relative risk is 1.029. Note that this for exposure at one month of age.

Now we are going to look at children whose exposure is at three months of age. So exposure at three months of age again is all cases where speech delay was seen at least twice. I'm sure you have all caught on, so I'm not going to belabor this, but you can see in fact that I think we can say the relative risk certainly does not disappear and doesn't vary much.

Now with autism, if we limit it to children with exposure at either one month or three months of age, and cases of autism that were seen at least twice, there is a relative risk that is no different than one and that is replicated whether we limit it to children with a diagnosis mentioned in the chart where the child was referred to a specialist, or the child was confirmed by a specialist. We see no difference from one. If we look at children where visits were more than twice and where the diagnosis was mentioned in the chart, referred to a specialist or confirmed by a specialist, I

don't see any evidence that there is a departure from a relative risk of one.

And now on to the final slide where we look at attention deficit disorder, attention hyperactivity disorder. Looking now at exposure of one month of age. If we look at all cases where they were seen for ADD at least twice, the relative risk is 1.006 with wide confidence intervals that include one.

Restricting it now to cases where the diagnosis was in fact mentioned in the chart, relative risk is still close to one. Referred to a specialist, relative risk of 1.007 and where a diagnosis of ADD was confirmed by a specialist, again 1.01 with confidence intervals somewhat wide. Today at least, and including one.

Where we look at exposure at three months of age, looking at all cases, relative risk of 1.008 and now with a confidence interval that skirts significance of 1.000, with an upper limit of 1.016. When the diagnosis is mentioned in the chart, it is about the same. When we limit it to children who are referred to a specialist, or confirmed by a specialist here in particular, the relative risk is 1.021 with confidence intervals now that exclude one.

One might say that these are eight relative risk calculations, however, they are certainly not independent, so I'm not sure that multiple testing actually holds in this particular case.

So I am going to wrap up. I'm not sure that we should actually have questions right now. Maybe one or two, but I think this would lead best right into Phil's discussion, unless there is some burning questions that simply can't wait.

Dr. Verstraeten: Just something you ought to mention. This condition of having been mentioned at least twice only applies to speech, not for ADD or autism.

Dr. Davis: Thank you, I did not understand that.

Dr. Cordero: Just a clarification on the autism, did you find in the record review any evidence of regression or was that possible to get out of the records?

Dr. Davis: There were only 13 cases of autism and I looked at a good number of those. I was actually looking for that out of curiosity. I don't recall any cases that I ran across, and I don't know if Frank or Tom, I don't think so. We had Chart abstractors do the review at NCK. Did you happen to hear about that? It's a very specific type of autism that it supposed to occur in about 20% of autism where a child is normal until some time of age and then has an acute regression.

Dr. Stein: Did you say the 1,000 cases you reviewed were randomly selected charts?

Dr. Davis: Were these a random selection? I'm trying to remember. Certainly everybody was speech delay, autism and attention deficit at Group Health. I am trying to remember if they were random selection at NCK.

Dr. Myers: I think they were all cases of autism, all cases of ADD and all cases of speech that were mentioned twice.

Dr. Sinks: You did a very nice job of looking at these records, and I want to complement you on that. It strikes me that what you're really showing is how well the records are reflected in your automated system, and not necessarily that these individuals are more or less true cases, because in fact except for the last one you showed, almost all of these

cases were in fact down this differential in terms of being referred and they are almost all the same set of kids.

Dr. Davis: Yes, I think in a previous lecture Tom talks about it this way and I have no better way to put it. He said something is apparently worrying these parents and they are bringing these kids in and that's causing them to show up repeatedly. Now whether that is a measure of parental over reaction, I don't think we can discern that, but they all seem to share the same attributes. So almost no matter how much you slice the pie, they all seem to be going through this data set with the same set of covariants and exposure metrics.

Dr. Myers: We're not defining a true case by a different set of diagnostic criteria other than the specialist has agreed with something else, but this is the case.

Dr. Johnson: We can have some more questions on this subject later, but let's let Dr. Rhodes do his thing.

Dr. Rhodes: Thank you for inviting me to speak today. First I want to commend Dr. Verstraeten on more work than I would ever do in the course of a couple of years.

I think it is important to understand, I have been looking at the data set for about one month and Tom and others have been looking at the data set for upwards of six months or so. I am not going to comment on everything he has done. Obviously some of the things he has done are quite new and I have not taken a look at those.

I think I had sort of two purposes in mind in going through the analyses I've done. One was just a very quick verification that there wasn't some crucial missing statement in 4,000 lines of programming, and there wasn't. Tom's programming was all perfectly clear.

I also wanted to try to take a different look at the data because I think some times we make choices soon in our analyses. We conceptualize the problem very quickly and then everything else kind of depends on those initial choices and we don't always go down other pathways.

I will take a few minutes to talk about what it is I think we are about in this data set. What questions are answerable in this kind of data. Where does Thimerosal into that continuum and I will talk about what I saw as at least some possible difficulties with Tom's early analyses, just in the sense that there were things that raised red flags with me and I know they would with other people. It doesn't mean that they would affect the analyses by taking into account, but that they were worthy of at least taking a look at.

I think we will see that I will approach the data analysis in somewhat of a different way, and I will talk about what some of the results are when I look at the data in somewhat of a different fashion.

The Vaccine Safety Datalink study data set is an amazing resource that is very good at doing certain things, and not so good at doing other things. In terms of vaccine safety it is good to excellent in evaluating exposure outcome pairs where the outcome is acute, medically well-defined, has a high probability of coming to attention and has a clear onset occurring a short time after exposure. Especially if the effect of the exposure on the outcome is transitory and this is still possible and works even better in cases where the exposures are almost universal, but there is some sufficient variation in the age and exposure. As an example in which the VSD is very good at finding an association for example is seizures occurring after DTP or MMR.

Now if MMR had the effect of raising the relative risk of seizures forever, it would be much more difficult to study.

Those pairs that are harder to evaluate is where the outcome is chronic or not so medically well-defined. For example, speech delay. Or where the onset is not well-defined, and in these cases if the exposure is nearly universal, we are really stuck with trying to compare groups that do or don't have the exposure. In many cases them, the group that doesn't have it will be a small, unrepresentative group. For example, if we are trying to study the effect of attention deficit disorder after MMR.

Now you might think I am going to say it is impossible to study Thimerosal in this cohort, but I am not going to say that. But where does Thimerosal in developmental delays fall in this continuum?

The outcomes here certainly do vary on their medical certainty. There is quite a bit of difference on autism versus speech delay in terms of medical certainty, and also the likelihood of coming to medical attention at some point. For example, just the orientation and the facilities available at the different HMOs can have a great effect in terms of whether certain things come to medical attention or not and/or are followed up in that context.

These outcomes in most cases are chronic and the time of onset is not well-defined. We are also in a situation where the exposures are nearly universal and others have argued that the completely unvaccinated do form an unrepresentative sub-group?

So are we in a hopeless situation? No, there is variation in the amount of Thimerosal by the type and manufacturer of vaccine. If there wasn't, or if there weren't changes in

vaccination policy over time, then we would be in a more or less hopeless situation.

People have also eluded to this. Are we studying differences in cumulative Thimerosal exposure at some age? Well, that is what we are studying, but are these differences in cumulative exposure due to the policy of the HMO or the clinic we are talking about, or due to the self-selection of the parent. For example, lateness in getting vaccinated, a reluctance to accept any vaccination or medical care?

Now, just as in the kitchen where the chef chooses the ingredients they are going to use, the kids you choose to let into your analysis can have a great effect on what happens eventually.

In one of the areas in seeing some of Tom's early presentations, I did have some concerns and I thought that others would have concerns. And even if it ultimately had no bearing on the outcomes, the fact that certain choices had been made might cause some problems. One of these was in the sense of what exclusion criteria was set in terms of the kids being the analysis.

To briefly summarize, they had to be born into the HMO. I have no problem with that. We are looking at early vaccination exposures at an early age. It is crucial that we feel that we have that exposure information, so I have absolutely no quarrel with that.

Follow continuously for at least one year, and he didn't really mention this, but that we actually only use their first follow up period because some substantial number of kids do dis-enroll and come back to the HMO. I have some problems with this, although not too much in some context and a little bit in other context.

Some of the others that will cause more concerns are that there is no using of prematurity codes, although in some cases they are almost synonymous with low birth rate codes. Probably one of the biggest is they not have one of the many possible perinatal conditions. A more minor one is that they not receive any hepatitis-B immunoglobulin and one that probably should be a little more controversial and hasn't been is whether they get two or more polio vaccines by age one.

I want to say at the beginning that all of these exclusions had good intent and good thought behind them. They weren't just randomly chosen exclusion criteria. For example, the prematurity exclusion. It is easy to see that these kids, certainly at the extreme values, would be much less likely to receive HepB and other vaccines, but especially HepB at an early age and they may be much more likely to have some of these outcomes of interest. So especially if we are looking at the analyses at one month, if we leave these kids in, we are going to put high risk kids into the unvaccinated group, unfairly raise the baseline rate and unfairly or at least miss an association if one is there.

Similarly, I think children received two or less polio vaccines in their first year may not be accessing the system as often as others or they may have a very different outlook on what constitutes a condition that requires medical care.

I think I am not the only one that has been struck by the difference in what has been caused by some of these exclusions. For example, there is a whole range where actually there were 23 separate ICD-9 codes that were included in the so-called perinatal exclusion codes. When you look at this by HMO, there is quite a difference in terms of how many kids get excluded from NCK versus Group Health. About 19% at NCK and about 7% at Group Health, which is certainly a startling difference.

Also, if you look across the different birth facilities at NCK, you see a range of about 13% to 36% of the kids are being excluded just on the basis of these codes. This doesn't include whether they are excluded by other codes, it is just looking at that possible exclusion criteria.

Now some of these ICD-9 codes are likely to represent fairly minor occurrences. For example, 767.1 is scalp injury at birth. They are also very different across the HMOs as you would expect. There are over 6,000 kids being excluded at NCK for this code, and only 24 at Group Health. 779.3 indicating some sort of feeding problem. About 4,000 at NCK and a little over 500 at Group Health.

The prematurity codes are also coded differentially at the two HMOs. About 5% at Northern California and a little less than 2% at Group Health. I don't know if you are familiar with this, but they actually do have a fifth digit that gives you some sense of what the birth weight was, and from these codes you can see that over a third of those that were excluded at Group Health actually are not low birth weight, but they are premature. Only 5% for those at Group Health, so obviously there is a very different style of coding for the prematurity code at those two HMOs.

The other exclusion criteria of interest, two or more polio vaccines in one year and if the first enrollment is greater than one year, obviously this has some concerns if you start wanting to use events that occur at less than one year. Some statisticians would take offense at having exclusion criteria that happen after the event. It is rarely wise to condition on the future. It's like counting your chickens before they hatch in some respects. Although certainly then if you only count events that happen after one year, there is no problem in doing that.

I think one problem that I have not found any solutions for, but there are substantial problems, at least at NCK, in terms of the enrollment dates. There are a lot of kids who come in and out of the HMO for various reasons. It is not clear. They may have three or four enrollment periods. They also actually receive a lot of care during these so-called disenrollment periods, so it is not clear whether their disenrollment is related to the fact of their parents changing jobs or whatever, they are still covered and it just appears that they are disenrolled. That is a problem I have not solved. One thing it does add up to, when you make all these exclusion criteria and you look at some of the outcomes, you will see that around half of the total events have been used in some of these categories, which is certainly of some concern.

Another area where I have had substantial concerns and I think others also, remembering back to Tom's slides about how many kids fall into the different exposure levels. You remember across the two HMOs combined in his cohort, there were about 2% to 3% of the kids were in each of the zero, 12.5 and 25, then a huge jump when you reached the other ones. So about 7% of the kids were in the zero to 25% group and over 90% in the other groups. I certainly had concerns that they were an odd group in some ways, and other people have also raised those concerns. So just some further evidence that they are not like the other kids.

For example, when you take the three month classification and say what happens to these kids a little later on? If you look at them even seven or 14 days later, you can see that there has been substantial movement from the zero and the 12.5% group. For example, after seven days at NCK, fully 27% of the zero group has received some sort of vaccination in the next seven days, and 42% have received some vaccination in the next 14 days. Some of those are

receiving 62.5 micrograms of Thimerosal in that 14 day period.

Now the 25% group is much more stable than the 12.5 at that point, and if you look at the 37.5 or 50 there is hardly any vaccination in those groups as you might expect. They have basically received what they are going to get until they reach the next milestone.

To a large extent here, at least in the zero group and to a large extent the 12.5% group, we're analyzing lateness more so. We are certainly analyzing a difference in Thimerosal burden by age, but if you move the line back a little bit to three and a half months, you would have substantially different exposure groups. At least in these lower exposure groups.

This one is a little busy, but it is very much in line with some of Tom's slides. That there may be less medical care utilization in the low exposure groups.

This is the average time since the last well-child visit (ICD-9 V20.*), from the dis-enrollment time back to the last well-child visit. How long has it been on average, and if you look across and this is how long you've been followed. For example, those who were followed greater than 48 months, there were 1,500 kids. This is actually NCK and not Group Health because there aren't this many kids in Group Health. So if you go across and look at 0-25, you can see that the average time since their last visit is somewhat longer for the lower age groups. It is not a very large difference, but for example, 19-24 months there is about a month or more difference on average since that last well-child visit. If you look at the proportion of those kids who have had a visit say within the last year or the last two years and you get to the older age groups, there is a

reasonable difference between those who are in the higher exposure groups than those in the lower groups.

Another factor that was raised by the CIs at NCK was that there can be substantial clinic differences in California. Northern California is geographically much more widely dispersed than Group Health. Group Health is essentially a much smaller area than Northern California.

And birth facilities and clinics often do have different policies. For example, the use of HepB vaccine in the first month of life, and this is for all children born into HMO in 1992-1998 at NCK, there was a range of 4% to 85% for any usage of HepB in the first month of life, with an overall mean for all those kids at about 43%.

There are great differences in the exposure groups. I haven't defined these yet, but we will see this in a moment. I through V range from 37.5 micrograms. II and III are different again, to 50 for 62.5 and 5 of 75, and I will define these in a minute. But just to show that for the largest clinics at NCK had very different distributions of those five exposure groups.

As my epidemiologist friends in the audience will point out, they vary on exposure, so we don't carry them unless they vary on the outcome. Well, the clinics do vary on the outcome, although of course at this point you don't know if they vary on the outcome because they vary on the exposure, but at least we can verify that there might be some chance for confounding at this point.

For example, taking the category of all developmental delays and looking it at by clinic and all children followed longer than four years, there was an overall percentage of these conditions at 4.4%. For the 32 clinics that any substantial number of kids, there was a range of 1.6 to

8.7%. This is the distribution of how many clinics had what percentages. So I think there was a reasonable enough variation there. Of course that variation could well be due to the exposure, but at least at this point I think we have enough evidence to think that clinic is worthy of consideration as a possible confounder.

Again, this information is available at this point only at Northern California. It is not available at Group Health.

I think at this point I was led to the idea and I sort of stepped back a bit. We have had the question posed of, can you answer the question of what is the effect of Thimerosal, going all the way from zero up to 25, up to 50, up to 75 and through 100. My various explorations through the data led me to basically think that some of these questions could be well answered and others could not be well answered from this data set. Those answers that I thought could not be well answered from this data set, were answers that involved questions of what happens between zero and something? What happens between 12.5 and something? But if you look at the data that is available and how those data occurred, some times nature conspires to take observational data and make it almost look like an experiment. Some times it doesn't. In this case I think the closest we can come to regarding this as an experiment as opposed to totally jumbled and meaningless observational data is to think again in terms of what exposure groups do kids fall into and how do they get there?

This is similar to some of the slides Tom showed before, but essentially when you look at the data there are five large groups the kids fall into. These totals are going to be between 85% and 90% of all the kids that have entered into Tom's analysis.

There are five Thimerosal levels ranging from 37.5 to 75. There are two ways to get to 50 here. When you see what happens here, it is also very helpful to see that all five of these groups at NCK, but only three of them occurred at Group Health, and that is an important thing to keep in mind. At Group Health, it is very balanced that over two-thirds of these kids fall into one of those groups. And when we start combining these two things, we can get funny types of analyses in the sense that we have to understand and remember at Group Health there are a few kids who have 37.5, but any comparison you are seeing of 37.5 to anything was coming from Northern California. It was not coming from Group Health. Similarly, analyses that were coming from 75, although there is an equal number of kids from those two HMOs, some of the event rates were so much higher in Group Health, that the 75 group was being dominated by Group Health as compared to NCK.

In these analyses you can get very different results when you throw these things together, as compared to when you make head to head comparisons of some of these groups.

So at this point my thinking was that if you want to talk about the effect of a difference of 25, at level of 50 versus 75 or 37.5 versus 62.5, this is a good data set to do it. These kids are achieving these levels, mostly based on policy of the HMO or clinic at the time they are getting vaccinated far more than they are on lateness or anything else.

Some comparisons kind of jump out at you in the sense that we certainly would like to compare the smallest group to the largest group. That is 37.5 is the biggest differential we have. Some of the comparisons are a little more natural in the sense that if you think back to two of these groups that differ on whether they receive a DTP-HIB combination or

whether they received separate DTP and Hib, and actually at this point they could have received a DTP and HIB separately or a DTaP and HIB separately. And there are even a few kids who would have received a DT and a HIB separately.

My approach was to think of terms of the analysis of the zero to 12.5 and the 25. I am not advocating totally throwing them away and never considering them in any analysis, but at least for now let's think if we can establish if there are differences in this group of 37 to 75, then in a sense we really don't need them. If we don't see any difference in this group of 37 to 75, it is not that we are home free and we feel everything is okay, but at least we started from a place where we feel the data set has good information to offer us. And if we are going to include those other groups, we are going to have to think very carefully about how we are actually going to do it.

In terms of how you would approach exclusion criteria in this study, I would have a fairly different point of view at one month and three months.

At one month there are still some problematic aspects to this. I am not going to try to base what happens to them at three months in terms of a one month exclusion really. At least it is not very satisfying to do that.

Here there is still some question about what are appropriate exclusion criteria at one month. I think most of the interest, at least in Tom's analysis, has been at three months.

There is an exclusion policy that just says the price of admission to the study is having achieved one of those exposures by three months, end of story. Don't tell me you had a code of 647.2 at seven days or whatever. If the

choice is made to give you those vaccines by the time you are 93 days, you get to be in the analysis.

The other showing the clinic was an important variable and led me to think it was important to think of the clinic as an additional stratification variable at NCK.

The one very sobering thing that has been eluded to and is not obvious from the analyses that Tom presents, is that when you stratify very finely at time of birth, again these five exposure groups are very largely a matter of policy. Policies change very quickly over the course of few months. What is not apparent is the effective sample size, the effective number of cases that enter these analyses are often very different than the total number of cases that you see quoted. I could actually work up some slides for tomorrow that show how many cases really do enter some of these analyses. You may start with 3,000, but I think in some cases you may be down to 300 in terms of cases that actually affect the analysis.

To try to wrap this up, if you were just looking at two exposure groups, for example the DTP-HIB combined versus the DTP and HIB separately, at NCK this policy choice is implemented and happens over the course of about two months. If you stratify finely enough and the policy changes are made quickly enough, you have no analyses because no one would be temporarily overlapped in order to be compared.

The other thing that happens at NCK is that even a year or two years after the policy change has been made and all kids are supposedly receiving the combination, there is an odd, small group of kids that supposedly receives separate DTP and Hib, and an unusually high percentage of those kids are outcomes.

Then when you go back and look at their data, there is supposed to be information on where they received their vaccinations, the manufacturer, lot number, et cetera. Typically for some kids the facility is missing, the manufacturer is missing and one suspects these kids are those whose charts have been missed or pulled for various reasons and there have been data quality issues with some of these kids.

For example, if 1,500 kids were receiving one vaccine combination in that month of birth and 20 were receiving some other, I have removed the 20 completely from the analyses. In essence, the right thing to do might be to put them in with the 1,500 but at least for now I have left them out.

So the question is not so much the choice of the five exposure groups. There were two themes that came up in a lot of Tom's slides. One was using the zero group as the comparison group and looking at how wide all the confidence intervals were for the other exposure groups, and did they or did they not overlap.

Well, the secret is you pick a different exposure group as a comparison, all those confidence intervals will be different and some will overlap and some will not. So that is really sort of a false issue in some way. Also, the number of events was always very small in that group.

The other thing was that his test for trend, which I philosophically don't like very much because they ascribe a difference of zero to 25 is the same as 25 to 50, as the same as 50 to 75. I think in the end when you have enough separation and you know that your data kind of looks like that, I think it's okay as a summary. But I have a philosophical problem with running with that analysis as sort of your major type of analysis.

He was also claiming though that if you left off the low groups, you could still see the trend in the groups 37 to 75, because that is where most of the cases are. So while I have left off the small groups and you will no longer see a comparison of 75 to zero, or 62 to zero, you still will see something that would tell you if there would be a trend in Tom's analyses or not. Again, I am still using the bulk of his data, at least in the initial analyses.

This is the only analysis I am going to present at one month. This is a combined analysis of NCK and Group Health. Using more or less Tom's original cohort and just saying any or none, and using the code 315.3*, we get a relative risk for the anti-Thimerosal group of about 1.2, chi-square 12.1 and various significant P-value.

Now adding a clinic, it doesn't drop the relative risk very much, but it does increase the variability quite a bit. Now some clinics have almost nobody with HepB at one months. Some clinics have 90%. At one month you might say there is over-stratification, but I think it is worth considering here.

Now I take all those kids that Tom has excluded based on prematurity exclusion codes and throw them in. At one month I think there is some argument that is overdoing it. Throwing them all back in. I think there is a clear argument that is going too far, but that further brings things down. I try to bring it back up by bringing in those premature kids who were less than 1750 grams. It brings it back up a little bit. Make the polio exclusion, it brings it back down. So you can push, I can pull.

But there has been substantial movement from this very highly significant result down to a fairly marginal result.

I think one could argue a long time what is the appropriate group to have at one month. Again, if you agree with my premise of what are these data good for, I think there is much less things to argue about at three months.

This is just presenting data from NCK for the moment. Looking at the reference group here is 37.5 micrograms, so we are comparing our four groups to that reference group. We are looking at all the developmental disorders, which is the largest group. Due to lack of room on the slide, I have presented just the relative rate and the P-value for that relative rate.

For example, if we start from Tom's original cohort, these are all elevated compared to 37.5. Two of them are significant at the .05 level, but not too far beyond. One is very close.

This one has a lot of cases, so it even has a very low relative rate. It still has a significant value.

You could put these in different orders, but as I go down the list here, everything I have done in two, I will have done in three. So I am adding or doing various things. For example here, these are including all the kids that Tom would have excluded for various reasons. One group actually goes up. This group is very close to one. This group goes up a little bit because a lot more cases are being included. This group comes down a little bit.

Now putting clinic, we see this one stays about the same. A lot of these come down quite a bit. The P-value are not becoming very impressive.

Now leaving out those kids that have the so-called odd codes, I looked at the pattern of exposure group based on birth month and if there was some category, for example if

you had category four and there were say 15 kids in a particular birth month that had that particular exposure pattern, I said most of those are probably due to coding problems. Let's leave those out and see what happens. That is something that should be followed up and verified that those indeed were coding problems.

At this point three of the groups are still a little elevated, but none of the P-values are lower than .2 at this point.

Now looking at the speech and language delay codes, 315.3*, instead of going through all the intermediate steps, I just do the original and then the final, which we would be using clinic. Putting back the excluded kids and tossing out the small number of kids that have odd codes. You see that they go from not quite significant, but fairly large relative risk to almost nothing. Are they significant? Nothing. High but not significant, nothing. Not much change there.

Remember, Group Health did not have all exposure groups. They only had groups two, four and five.

Here the reference group is 50 rather than 37.5 because that is what Group Health has. Here there was not much going on before. Maybe even a little bit more going on in this group afterwards, but very little change there.

We have added more cases, so the P-value is a little bit lower.

Then putting together Group Health and NCK, but just using Groups II, IV and V. It actually will change things if we include Groups I and III, but to avoid that for the moment let's just focus on Groups II, IV and V. Putting them all together. Not much going on before. Not much going on after.

So what have I concluded from my reanalysis? You don't have to agree with me, but these are my conclusions.

That there are strong uncertainties about the fairness or the comparability of the low exposure groups.

That these concerns are much less for the groups starting 37.5. But that the evaluation, even of these groups 37.5 to 75, is still somewhat tricky because of several issues.

That the small amount of calendar overlap for the use of these different policies and for the policies that led to the various exposure groups can really affect our analyses in two ways. One that very many of the cases totally drop out of the analysis, and that some cases who have been maybe miscoded can actually have a very undue influence on the result.

If we have 1,400 kids in one exposure group, 10 who are miscoded, and maybe their miscoding is also related to the fact that they are a case and it did actually occur, in some of these birth cohorts you would see three cases out of 10 kids, or a similar number out of 1,400 kids, it is clear those kids are having an undue influence on the results.

I think it is clear that at least in some respects the original exclusion criteria were too extreme. I don't think they have affected things as much as, for example, accounting for clinic practices at NCK, but I think it was worth taking that step of thinking what were exclusion practices that wouldn't at least have caused people to have trouble with the results.

Overall there were still some slight tendencies for the higher exposure groups to have somewhat higher rates, but the P-values were in general quite unimpressive and for the most part were .20 or even much higher.

What do I perceive as being some of the limitations and/or extensions of what I have presented here? I think it is reasonable to argue that a complete rejection of these low exposure groups may be too severe. The 25 group may not be nearly as bad as the 0 and 12.5 group and one may be able to do something with that group.

I think one cannot certainly take data where there is such a restricted range of Thimerosal and say Thimerosal is fine, give as much as you want. We looked at a restricted range of Thimerosal in just one particular way.

I don't see any big differences in these groups looking at this. That doesn't answer all possible questions about Thimerosal. If you don't have those ranges, you can't answer about ranges you don't observe in your study.

I think I would say that I don't feel there is any fair way to compare 0 with say 50 or 75 at three months, at least in the data as we now have it.

I used a fairly crude measure of clinic at NCK. I think with a little more work one could use a better measure and actually track his over time. Most kids do stay at the same clinic, some do change. I just picked where did you go most often, but obviously there are changes in where you go and that could affect things. But with a little more coding and a little more time, one could actually track that a little better.

I think it is very important to check the assumption that these kids who have these unusual coding patterns at NCK are actually in fact miscoded.

But I think what it also argues for is that in fact the data were too stratified by month of birth and that there should

be some backing off, so these temporal overlaps don't throw most of the cases out of the analysis.

I will now welcome comments or private discussions with Dr. Walker and Dr. Oakes on what might be a proper way to accomplish that. A fair way.

Certainly as already has been evidenced, the data from the chart reviews have been used to refine case definitions, at least in the analysis that has been done so far. They certainly haven't made things go away in those analyses.

Certainly there are also even from these two HMOs, there is more variability in the exposure within birth cohort in the latter part of the follow up. In '97 and '98, there was much more difference in terms of exposure categories, so as these cohorts age you have more an opportunity, at least in these restricted ranges.

What would one want? One would want somehow within a situation where there is comparable ascertainment, you would like kids who got very low level, 0 to 25, whatever micrograms of Thimerosal, versus those that got 75 or 100, but to have the same number of vaccinations. That they are equally vaccinated, but because of policy differences or manufacturer differences, have big difference in Thimerosal usage. But also in which you have equal ascertainment, and that may be rough. Getting both of those at the same time may be very, very difficult.

Dr. Johnson:

We have a fair amount of time dedicated at the end of this day for debate and rehashing of the data presentations. I think it would be better to move on to biologic plausibility than take a break, and then come back and put all this in one pot and discuss it. If you have a short procedural question, Dr. Guess, that will be fine.

Dr. Guess: Can we get copies of those overheads? It is very difficult to take notes. That was really excellent.

Dr. Oakes: Because there was a lot of information for those who haven't seen it before.

Dr. Bernier: We will try to do that, but we failed to put on Dr. Rhodes' graphs that there may be errors of fact or omission, and for our purposes this is a very important piece of writing to put on every one of those, so I am reluctant to release that until we try to get that done tonight and have these for you tomorrow. If you need one to look over and you can turn back over to us tonight and promise not to copy, we can maybe we can do that. But I think it would be best if you get them tomorrow when it has that on there.

While I am making that point, let me just reemphasize if I could the importance of trying to protect the information that we have been talking about. As many of you know, we are invited here. We have asked you to keep this information confidential. We do have a plan for discussing these data at the upcoming meeting of the Advisory Committee on Immunization Practices on June 21 and June 22. At that time CDC plans to make a public release of this information, so I think it would serve all of our interests best if we could continue to consider these data. The ACIP work group will be considering also. If we could consider these data in a certain protected environment. So we are asking people who have done a great job protecting this information up until now, to continue to do that until the time of the ACIP meeting. So to basically consider this embargoed information. That would help all of us to use the machinery that we have in place for considering these data and for arriving at policy recommendations.

Dr. Johnson: Dr. Koller?

Dr. Koller:

You are probably wondering why a veterinarian was invited to address this distinguished group of professionals. That question is not for me to answer, but I am very pleased to have been invited to participate in this meeting and to enjoy the beautiful surroundings of this facility.

Most of you do not know me and I do not know you, so I thought I would give you a brief background for myself.

My background has been quite diversified. I am a D.V.M., Ph.D. primarily in pathology, but my research was with oncogenic viruses and immunology. I then took my first job with NIEHS, where I pioneered the field now known as immunotoxicology, then quickly moved to academia and had worked many years in my field of interest, which is pathology, toxicology, immunology and carcinogenesis. Evaluating the effects on numerous chemicals, including mercury and methylmercury, and today I am presently focusing primarily on auto-immune diseases.

It is interesting that I have a publication here that came out in the year 2000 of the Journal of Auto-Immunity. The Title "Vaccination and Auto-Immunity". Vaccinosis, a dangerous liaison. So this is another aspect of vaccination that is of concern to the medical professions.

I was even foolish enough to venture into administration. I was Dean of the College of Veterinary Medicine for 10 years. Then I have moved back into a more relaxed, rewarding life of a professor in the same college.

I have served on many national committees, mainly for EPA, ATSDR, NCI, National Research Council, Institute of Medicine, National Advisory Committee to establish acute exposure guidelines for humans. Most all of these focusing on establishing standards for human exposures.

I am also presently involved with the Army and CDC in establishing human guidelines for nerve agents. As you know, we are trying to destroy all those stockpiles.

I want to start with a disclaimer. When Roger called me I was just finishing up some reports for the Institute of Medicine committee regarding an update of health effects for Vietnam veterans and was starting to prepare for a grant renewal. I quickly dropped that and did a rush review of the toxicity of mercury. Primarily methylmercury, so you will have to pardon if I am not as thorough as you would like to see, particularly on some of the basic mechanisms.

First slide please. Most of you are familiar with the neurologic symptoms of methylmercury. There are many of them. Tremors, emotional lability, insomnia, memory loss, you can see neuromuscular effects, headaches. Pretty common of a lot of things. Polyneuropathies, several of them. Performance deficits have been recognized. Hearing and visual loss. Even hallucinations and photophobia.

Next slide. What I want to do was show the daily consumption of methylmercury, and it might surprise some of you. For infants six to 11 months of age, about .5 micrograms per day. Two year olds, 1.3. Females 25 to 30 years, around 3. Males, 3.9.

On a body weight basis for the intake, it is equivalent to about 0.05 micrograms per kilogram per day, except two year olds and that would be a little higher.

For health professionals the values are higher. 8.2 for females, 8.6 for males. Health professionals probably eat more health and eat more fish.

Canadians also consume a lot of fish, so you can see the values are higher.

The FDA estimates the average intake of total mercury to be somewhere between 50 and 100 micrograms per kilogram per day.

Now the ATSDR establishes their minimum values on a study. It is a Seychelles child development study by Tom Clarkson and his group. It is somewhat of an ideal study. They have 700 mother/infant pairs tested from parturition through 66 months of age. Actually it's before parturition.

Mercury levels are about 10 to 20 times higher than in the U.S. due to the consumption of fish in their diet. The environment is quite pristine. The population is high literate. They are quite healthy with low alcohol and tobacco use.

The developing fetus was exposed in utero, which as we go through some of the data today is going to be extremely important because we know the developing neurologic system is more sensitive than one that is fully developed.

Neonates were continued to be exposed via breast feeding.

What is interesting is the relationship of mercury in the blood, or in this case the hair, of the mothers versus the children. They are pretty close, and I would assume that even though this was at 66 months of age, the 6.5 ppm, that would probably be very similar as an infant and a newborn. Particularly because methylmercury can cross the placental barrier.

Six Neurobehavioral Tests were conducted on children at 66 months of age. Quoting the articles, "none of the tests indicated an adverse effect on methylmercury exposure" and in fact, "four of the six measures showed better scores in the highest methylmercury exposed group."

Remember before I remove this, that even though this is the mean, this is the range. So we are looking at some of these individuals had quite high levels of mercury in their hair.

The six Neurobehavioral Tests, and I am sure most of you recognize them, were the General Cognitive Index of the McCarthy Scales of Children's Abilities.

The Preschool Language Scale total score.

Letter and Word Recognition.

Applied Problems sub-tests of the Woodcock-Johnson Tests of Achievement.

The Bender Gestalt Test and the Total T score from the Child Behavior Checklist.

These are backed up. I grabbed these as I left town. That's Faeroe Islands. There were other studies that have been considered to establish standards. One is the Faeroe Islands where 917 children seven years of age were tested. Basically their conclusions are that the neuropsychological testing indicated mercury related dysfunction of language, attention, memory and visuospatial and motor function remained. That means they still saw these after children and women with maternal hair mercury above 10 ppm were excluded.

The problem with these studies is there were several confounding factors. There were higher PCB levels in these individuals and there were other factors. So it's not as pristine an environment as you would find in the Seychelles population.

In another group, the Amazon River Basin, 91 adults with hair mercury less than 50 ppm. Although the clinical examinations were normal, those individuals in the highest exposures to mercury had some restriction visual fields and displayed some disorganized movements.

Another study, Mancora, Peru, 131 infant-mother pairs. Maternal hair 8.3 ppm. Somewhat similar to the Seychelles. They found no neurodevelopmental abnormalities in children.

Well how about blood? The ratio of hair to blood generally is recognized to be around 250. I have seen publications anywhere from 140 to 416, but 250 is usually accepted.

The other thing that has not been mentioned here today that has to be considered is the half life of mercury in the blood, particularly the organic mercuries. That ranges from 30 to 90 days. The average is considered to be around 50 days, so one-half of the mercury will be eliminated in 50 days from the body.

Usually the hair values lag blood by about four weeks.

In the Seychelles study, the highest group had an average of 15.3 ppm mercury in the hair. That translates using a 250 ratio to about .06 milligrams per liter of blood, which is 61 micrograms.

Daily intake we won't worry about.

If you look at 6.8 ppm, the amount in the blood was .027.

Thanks to Dr. Clarkson, he gave me some data in a 2000 publication that came out in the Journal of Pediatrics and gave me more information.

There is an article by Stajich et al where he looked at children that were born to term. Took newborns before vaccination and discovered they had .09 microgram per liter mercury in their blood, vaccinated them with hepatitis, so it would be 12.5 microgram. Forty-eight to 72 hours post-vaccination, their blood levels were 2.24. That was their mean. The range was not very large. So if you take that, recognizing that this is a background, very low level, I did some rough calculations. If it's a linear arrangement, if 12.5 in a vaccine resulted in 2.25 in the blood, 25 would equal about 5.5, 50 to about 11 micrograms per liter.

Is it cumulative? Everything we've heard today is that we're looking at cumulative exposures. I would assume they would need to really model out your doses and model into it a half life, so it is not necessarily cumulative, but actually the blood levels would depend on the time between vaccinations or the intervals.

So I took and compared this data to the Seychelles. Recognizing that the mother's hair was 6.8, their daily intake 34 micrograms, that blood equivalent would be 27 micrograms per liter. That's calculated.

Recognizing also that this is a continuous exposure, not only as a child but in utero, so these children were exposed to this level of mercury in utero, as a neonate and during their childhood when they were breast fed. So we are looking at an equivalent in the children's hair of 6.5, very similar to the mothers, which would be approximately about 25 micrograms per liter in their blood continuously.

So I guess I can leave the final analysis up to each one in this room, as we have children with this level probably much, much higher because some of the children or some of the mother's maternal hair and some of the children were as high as 25, or probably four times higher than that

without any abnormal neurological functional signs on the cognitive tests that were run by that group.

Any questions or would you like to hold them until after break?

Dr. Johnson: Why don't we have a few short procedural an then we will have a chance after the break to probe.

Dr. Modlin: Just a question about your analysis here. This is a term baby. I assume this is a term baby at the fiftieth percent for birth weight. What if you did the same analysis for either a pre-term baby or even more importantly a term baby that was at the fifth percentile for their birth weight?

Dr. Koller: Well, the pre-term started out at .79 and ended up at 7.36. Much higher. But today the data we were considered about was term data. That's why I did not include it.

Dr. Orenstein: You mentioned in the Faeroes that at seven years there were some pick ups of language problems, attention, memory and visuospatial and motor. You also mentioned that there potentially were some other confounding issues. Can you make the same sort of calculations for exposure in the Faeroes, in terms of what levels might the kids have been exposed to? And do we know if any of those exposures were clinically significant? In other words, were these kids just picked up on testing or had there been any clinical attention because of speech delay or some other clinical symptomatology?

Dr. Koller: I have that paper with me, but I can imagine Dr. Clarkson could probably answer that question.

Dr. Clarkson: The Faeroes is a perspective study. There were no clinical effects whatsoever. They are simply based on an

examination of these children at seven years of age with a whole variety of neurobehavioral tests.

Dr. Orenstein: Blood level correlates and all. Did they attempt to look at...

Dr. Clarkson: Yes, the hair levels and blood levels. The correlation that they found in the Faeroes with the blood level and cord blood, versus the outcome of these tests at seven years of age.

Dr. Orenstein: And was it in the same level though as the Seychelles? How high was the core blood...

Dr. Clarkson: Actually slightly lower. Their average levels were somewhat lower than the average in the Seychelles for mercury.

Dr. Sinks: Just to point out, I think many of your assumptions still here underlie this basic premise that methylmercury and ethylmercury are similar in terms of the toxicology. I want to ask one question of Dr. Clarkson, because I have heard of this study, however you pronounce it. I have not read it, but I was wondering if you could comment if you have seen it; what you think of the quality of their exposure assessment was? I know your lab is very well qualified for looking at mercury and we have frequently seen problems in mercury analysis from a variety of places.

Dr. Clarkson: Which study?

Dr. Sinks: The one that actually is referred to up here that was published this year from Mercer University and I think Emory may have had a role in it.

Dr. Clarkson: I would have to look at the reprint again.

Dr. Johnson: Would it be possible to photocopy that General Pediatrics article?

Dr. Koller: Yes, he has it and I photocopied it for me. To answer your question, I have always considered the neurological effects of ethylmercury and methylmercury to be somewhat similar at a similar dose.

Now ethylmercury has thought to cause maybe some of the other organ abnormalities. Maybe more so than methylmercury, but I have considered the responses, the toxic effects to the nervous system to be similar at a similar dose.

And to answer the aluminum question from my point of view, I have worked with a lot of metals. The mechanisms between aluminum and organic mercury are completely different and I would not expect a synergism.

Dr. Sinks: Just one other comment. I do think it is important to weigh the difference between the quality of the exposure assessment, which was done in Dr. Clarkson's study in the Seychelles and the amount of history he has in terms of that, with one study that has looked at a small number of infants here and how much reliability we can place on that data.

Dr. Koller: Exactly. I think though if you would calculate back, and that is what I was attempting to do, calculate back from the Seychelles and the background on the human population, you won't come too far off from this right here. For an infant. For older people it is going to be much higher.

Dr. Weil: I have just a technical question. On your data from the daily intake of 34 micrograms per day, you assumed a blood level of 27 micrograms per liters.

Dr. Koller: I think that's...

Dr. Weil: That doesn't fit with the data...

Dr. Koller: Just a minute, I probably should say milligrams per day. Let me see. That's milligram per day intake.

Dr. Pless: Just a quick comment on the study. They were looking at premature infants and they had 15 of them and the confidence intervals, the range of measurements was extremely wide. It is hard to know how they sampled this little kids, and that is why perhaps they got such an incredible range after a dose of vaccine. And I think the measurements were done within 48 hours or about 48 to 72 hours after vaccination. They only had five term infants in that group.

Dr. Brent: It was a very interesting presentation. It's nice to have some data to discuss. You inferred that it was probably based on the half life, not cumulatively. That's an extrapolation or hypothesis or do you have some confidence in that?

Dr. Koller: What I am assuming is that if a child is vaccinated as an infant with 12.5 micrograms of mercury, by 50 days that is going to be half that value. So to be re-vaccinated in 60 days with 25 micrograms, the total is not going to 37.5, See, 12...

Dr. Brent: I understand. The other thing is with some biological of some chemicals, the more you are exposed to them some times enzymes change with regard to excretion and metabolism. Is that known for mercury at all or is it totally unrelated to experience with the substance?

Dr. Koller: I'd say Tom is ready to answer that one.

Dr. Clarkson: As you know, methylmercury and ethylmercury are slowly metabolized to inorganic mercury. The common mercury bond is broken. It's achieved in two ways. The microflora in the intestinal tract break down methyl to inorganic and that is how we get rid of it. Methylmercury goes through an entropathic recirculation from liver to bile, to intestine and back reabsorbed again and but for these obliging micro organisms in the GI tract, we wouldn't really get rid of it. So does the microflora break it down to inorganic, which is not well absorbed and comes out in the feces.

The other way it is metabolized is by phagocytic cells in almost every tissue in the body, probably including microglia in the brain. These phagocytic cells will also break down methylmercury. We don't know for ethyl, but it's probably the same mechanism. So to what extent these change would do us, it's not known. It's an interesting question, but that's not know.

Dr. Brent: Are we going to get a copy of that, too. It would be nice to have to read tonight so we could...

Dr. Johnson: Yes, there will be copies.

Dr. Koller: Incidentally, these values are correct and that's very interesting. I just went back and looked. It's .034 in the Seychelles. They are taking in about .034 per day and this is their blood level, so there apparently is an equilibrium at some point from the intake and from the excretion. So those R values are correct.

Dr. Weil: It's hard to reconcile that with the ethylmercury levels you have above.

Dr. Koller: Well, you have to remember, this is a continuous exposure all the way through. This is a one time exposure.

Dr. White: Are there experiments, particularly with rodents, in which the effect at different developmental stages were studied of the same amount of mercury per weight? Grams of weight?

Dr. Koller: Per gram weight. Not that I am aware of.

Dr. White: I mean a comparable dose by weight at different developmental stages?

Dr. Koller: I'm sure there might be, but I am not aware of one.

Dr. Brent: With ethylmercury, but not with methylmercury. There are a lot of studies with methylmercury, but not with ethylmercury.

Dr. Johnson: What are they with methylmercury?

Dr. Brent: Well, with methylmercury, the problem is you have the epidemic in Japan and then the problem with the contaminated wheat in Iraq, where you had severe neurological deficits, but the dose that those people received was massive.

Dr. Johnson: I am thinking about a careful comparative to...

Dr. Brent: Well, the animal studies, yes, and the animal studies of the rat have a threshold in the low milligram per kilogram from what I recall.

Dr. Johnson: Is it different at a different developmental stage?

Dr. Brent: Yes, in utero the embryo is most sensitive. Especially in the rat, the brain development equivalent to the human development is actually postnatal. A lot of people don't realize that, but the first week after birth the rat brain is equivalent to the mid-gestation brain in the human. That is

the most sensitive time. The major effect in Japan was reduced cerebellum and severe microcephaly and spasticity and severe mental retardation, but a very high dose.

Dr. Johnson:

Thank you, Dr. Koller. Perhaps this is a good time to take a break. We have about a 15 minute break allocated, then we will come back for discussion.

This time is now dedicated for open discussion. There were a lot of points raised in the early part of the day that I don't think we reached any kind of satisfactory endpoint on, and I am sure there are questions for the presenters. So this is the time for an open exchange.

Bob Chen and others are ready to show again any of the material that was shown this morning. Dr. Walker?

Dr. Walker:

This question is for Dr. Rhodes, whose analysis was very impressive and like a lot of people I find myself ticking off things that I was going to say as he covered material.

I was both pleased and concerned though as I looked at the clinic analysis. As you pointed out, by restricting at the clinic level and maintaining the time matching, that the number of informative sets must have been vanishingly small. That raised the variance and you suggest I think reasonably that we could loosen up the matching account for time in some way else. But it didn't really explain to the big effect the clinic matching had on the point estimate, and I am wondering if given the very small number of informative sets, if it wasn't a good chance that the difference was just statistical artifact and that we shouldn't extrapolate too far from the analysis you've present us so far?

Dr. Rhodes:

Could you say the last part again?

Dr. Walker:

The concern is that there was so few informative sets in the clinic analysis as it's been performed so far. We saw the point estimate go down and everybody was gratified by that, but I am a little concerned about jumping on an attractive result which was based on a terribly small number of informative sets.

Dr. Rhodes:

I think it would be good to quantify, and you make a valid point. I think it would be good to quantify both in the initial analyses truly how many cases are taking part in the analysis. Also as I made the point, I think there are some other sets of unfortunate influential cases that shouldn't be in the analysis, so I think it's good to quantify both those things. And also I think your point is valid, too, once we stratify in clinic to show how many additional cases have fallen out.

I mean one way to get them back in is to loose the boundaries on stratification. Actually I have done that to some extent and it brings effects down even more dramatically than other things I've shown, which I was a little hesitant to show them because what I have been taught as a statistician is stratify as finely as possible and if you back off and your effects change, then that was an example of confounding.

But I think in this case I am not quite so sure. I think the small number of sets may be so fragile that backing off is actually the right thing to do.

Dr. Wallace:

I think you could represent time richly with lots of nots and get all the advantage of the matching. One other piece since I have the microphone. I heard in the discussion some kind of equivalence assumed between an analysis of cumulative dose and a implicit requirement that there be a physical accumulation of the metal in the body for that to be an appropriate analysis, and I should say that we do

accumulative dose analyses all the time on toxic drugs that are metabolized and don't accumulate in the body. I don't think there is a necessary connection, or that the lack of accumulation invalidates the kind of cumulative dose analysis that's been done.

Dr. Johnson:

I think the point was made also this morning that a series of acute exposures are also picked up with that endpoint. Dr. Brent.

Dr. Brent:

I have been glad to be here where there are so many statisticians and epidemiologists because I need to learn a lot in that area, but I want to tell you about our field of birth defects.

One of the problems we have is there are about 60 birth defects and anybody who does a large epidemiological study looks for correlations with a particular environmental agent just on a statistical basis will end up with three birth defects that are statistically associated with that environmental agent, just on the basis of probability. And I noticed that in the table on page 9, you have that flow sheet that has all the correlations or relative risks that have been calculated. There are about 80 of them and about nine of them are positive. So some of them are there because you would expect them to be positive, just on a statistical basis. In fact one of them shows a negative association under neurological degenerative disease. The .987 with the relative risk both from the '95 confidence levels below one shows a negative association. How do you look at this data and which ones do you assume are the statistical ones and which one is the real statistical association due to an association?

I bring it up because we have had some major tragedies with statistical associations. The one that I can think of is the collaborative perinatal project. The collaborative

perinatal project was a 50,000 patient study from 1957 to 1965 where they looked at women who they registered in pregnancy and they looked at everything. They followed the children after birth to age seven and did complete neurological evaluations, IQs and as much as they could do in those years. When they got to birth defects, there was one association that came up. Congenital heart disease, and it was a very famous paper by Dr. Hynanen. I can tell you there were probably 3,000 lawsuits about the fact that progestational agents was associated with congenital heart disease. It took 19 years to remove that warning. The FDA in 1999 finally removed the warning on congenital heart disease after millions and millions of dollars of lawsuits and aggravation about a statistical association.

I just want a perspective from the statisticians and the epidemiologists as to when you look at data like this, what concerns them? How do they look at this and how can they explain to me what it means to them? Because it confuses me when I look at the birth defect data.

Dr. Davis:

Actually we were aware of the Faeroe Island the Seychelles Island data, so believe it or not you'll just have to understand that when we went into this particular study, my thinking at the time was fairly rudimentary. It was Thimerosal equals mercury. The stuff I knew from the Seychelles and Faeroe Islands knew that the primary areas of focus were going to be on language and speech development and believe it or not, we also knew that autism would come into play eventually. So we had to at least study that.

So I kind of view this as both hypothesis testing in some way, in a sense that our pre-study hypothesis was to look at language and speech developmental disorders. And I will be quite honest with you, almost everything else was to some extent a screening analysis. Except nephrologic

damage, because we did know that mercury in some sense has been associated with kidney damage in the past. So that was kind of our take on it on really the very first phone call. Frank or Tom, do you want to expand on it? In essence there were two components. This hypothesis testing/screening component as well.

Dr. Verstraeten:

If I can say something about the number of analysis. This question has been raised before. If you do 100 analyses, with a statistical significance level of .05, you will find five significant ones. That makes sense. However, the level of significance of some of these findings go to .001. So even if you do whatever adjustment, that would still be significant for some of these.

Now I agree that fortunately I didn't put the level of significance for these. That would have been helpful and it would be useful also to have that adjustment, but that is one issue.

The other one is that some of these findings are in a way consistent. If it would be purely random, then we'd pop up in all different places and not in certain patterns and what I think you are seeing here. So I think those are arguments that why I am not very worried about the multiple comparison issue.

Dr. Brent:

Dr. Jones brought up a suggestion when we were talking in the coffee break. The collaborative perinatal project had 50,000 parents. They registered them right from the beginning of pregnancy and then they followed them very closely. It was subsidized. Probably all those children had DTP. Was mercury in the DTP in the fifties and sixties? Well, that is still on computer and available to you. One of the things I have been taught about Epidemiology is repetition. In other words, if you could get another body of

patients and demonstrate the same thing, it makes it more convincing.

Dr. Verstraeten: I would be the first person to try and analyze that. I have been asking all over if there is another data set I could look at and try to replicate it in a very oriented manner without doing another analysis.

Dr. Brent: Well, it's on the eleventh floor of the Archives Building in Washington D.C. and certainly any government employee would have accessibility to that data.

Dr. Verstraeten: I just heard that we have those data here, so we can...

Dr. Modlin: I think it would be wonderful to analyze that data set. I would just remind people that the only Thimerosal containing vaccine that children got during that era was DTP and that they all got it, so that it may be that your opportunities for using that to analyze what we would like to see may be limited on that basis.

Dr. Verstraeten: Can I add another issue? Along the line of what you mentioned, what we are waiting for or what we are trying to do is replicate it in a different data set. There is a possibility, if I can mention this right now, that the people at Harvard Pilgrim are going to try and replicate our findings in their data set. One of the advantages is they have actually weight of all the children, so we will be able to do this by weight exactly. So we want to avoid doing multiple comparisons just for the specific outcomes that we are interested in. That's one and then at the same time at the U.K., there is another data set of the General Practitioners, where we have also asked them if they can replicate our findings there. So we are waiting for those results.

Dr. Guess:

I think that it is an excellent idea to replicate and those look like good places to replicate it. I would add one additional note of caution, however, with any replication and that is something that has been said before. That there are many people who haven't been diagnosed with speech delays who have them. Relatively subtle levels of speech delays, so that these data sets would still have the problem of potentially incomplete cases ascertainment, where the incompleteness conceivably could still be linked to the vaccinations. So that if the replication doesn't confirm the findings, I would feel a little bit better about it, but if the replication confirms the findings, you still have this problem with incomplete case ascertainment, possibly differential. So I think one way to get around that might be gradations of severity of the speech delay or things like that which might be harder and more completely ascertained. I am sure a lot of people have thoughts on that.

Dr. Davis:

Let me just carry on, we actually have a whole hour tomorrow to talk about sort of our future research strategy and talk about different studies that might be done elsewhere, including whether or not the potential sites have problems like you are talking about. The sample size or just whether or not they would be suitable to answer various of the biases that we have been worried about today.

Dr. Stein:

The correctness of this association of Thimerosal with neurobehavioral or neurodevelopmental problems it seems to be really depends on the quality of the diagnosis. That's your endpoint, so I have a few questions.

With the ADHD diagnosis, you looked really at the billing codes. You did your data analysis on the billing codes and yet it was only a little over 30% that had a confirmed diagnosis. Why didn't you do the analysis on those that

had a confirmed diagnosis of ADHD, rather than those that were just put into the billing code?

Dr. Davis: We did. Are you saying why don't we redo the analysis limited just to those children confirmed with ADD?

Dr. Stein: Yes.

Dr. Davis: We do have that. Let's just see if we can pull it up real quick. These are the results of the analysis. The ADD analysis, when it is limited to children who are seen at least twice. Excuse me, looking at the exposure. The exposure is calculated by micrograms of mercury received at one month of age. These are now limited to the number of children seen at least twice and then these are level I, II and III, let's see if I can call on my memory here. Feel free to come in and help if anybody remembers these off the top of their head. These are children who were at least seen by the practitioner, referred to somebody else and then confirmed. So these are in fact the relative risks for children who are confirmed to be ADD children, and you can see that the relative risks are almost identical.

If we look at the assessment of the analysis when we are looking at exposure at three months of age, now these are all children who were seen at least twice and then when we look at children where the diagnosis was confirmed by this specialist or referral on a referral, we see that's now statistically significant with confidence intervals that exclude one.

Dr. Stein: Do you have the same data for speech delay?

Dr. Davis: Yes, we do.

Dr. Stein: Is it a similar result?

Dr. Davis:

Yes.

Dr. Stein:

I think with ADHD, and this is a clinical observation, if you go to a specialist as I think they have both at Group Health and Kaiser of Northern California, that diagnosis is fairly standardized. Developmental behavioral pediatrician or a neurologist, I think we can believe in that diagnosis.

On the other hand, speech and language delay include a wide variety of diagnoses. I really have never seen a child sent to a speech pathologist without coming back with some diagnosis. It may be very mild apraxia or very mild articulation deficit or very mild expressive language delay, but I have never seen one not come back. Now you had some that didn't, I know, but I have never seen that.

I think you really need to look at what the diagnoses were and if they were significant, rather than just a diagnosis of a speech delay. That's very fuzzy.

Dr. Davis:

That is a very good point and it just proved to be impossible, on both a preliminary look and on our large look. We can wish for standardization in medical records, and I'll show you it simply doesn't exist. Some kids got Woodcock and some kids got Bailey. It's just a whole commish of stuff done to children and it seems to vary by who their referral practitioner is and the age of the child, and other seemingly random events. I think what you are asking for is wonderful, but I don't think we are going to get it just by looking at medical records.

Dr. Sinks:

I would just like to compliment the investigators for actually diving into the data that were available as much as they did. I think they have done a fantastic job of doing it. At the same time I think it is important that we realize this is not the same thing as taking these children and putting them through a standardized battery of tests to determine

exactly what they have. The fact that we see such discrepancies in terms of the proportion of these kids were referred, 30% or 40% for ADHD versus 90% for speech delay suggests to me that there is a very large difference in terms of clinical practice and referral patterns, as well as willingness to accept diagnoses from a referral physician. I think we can say that we can cull these things down, but whether or not they have a specific disorder relative to each other, it really does require more of a standardized clinical battery. I think at the same time they have done a great job of doing what they could with the data.

Dr. Davis:

Well, thanks. I think actually we all agree with you on that point.

Dr. Johnson:

Could I ask, Tom, does it help you at all, the fact that as you tracked it more and more toward the more precise diagnosis, the relationship held. The relative risk stayed more or less the same as you got closer and closer to the specialists?

Dr. Sinks:

Personally not so much in the speech and the first two things because there really wasn't much difference in the numbers. Almost all of these kids were referred. Almost all of the referrals ended up being a confirmation, and it didn't suggest to me that there was really much difference in those groups if I remember.

The more troubling one to interpret was the ADHD. The thing there that is troubling to me is why is it that 60% of 50% of these kids are not getting referred? Is it an issue of the degree of their condition or that the primary treating person is very comfortable in treating this case and it isn't just simply not referring them. So I am not sure. It is reassuring that there is something going on in the data. If it just disappeared, yes, but again I think in the first two

examples we are really dealing with almost the same cases. I don't think you are culling that many out.

Dr. Verstraeten:

Can I quickly comment on that? In the first place on speech, we only looked at those kids that were diagnosed at least twice, and we already saw that the relative risk was higher in that group than the general group. So already we started out with a group of children that were more likely to be truly affected than others. More ADHD, so that might explain why the percentage was so high for speech.

The other issue, I think you are right about the ADHD. I think there are a lot of General Practitioners or Pediatricians who feel comfortable treating these kids straight away with Ritalin or whatever it is.

Another remark I would like to make, I don't think that this one completely takes away the concern about parental bias, because one could say the more concerned the parent, the more likely they will see a specialist and the more likely the specialist will treat the kid if the parent really insists.

Dr. Rapin:

Regarding the language disorder, you must have made on age when this diagnosis was made, yes?

Dr. Verstraeten:

Yes.

Dr. Rapin:

Okay, two things. Number one, perspective studies have shown that a large number of children with early language delays diagnosed say at two years or three years, by the age of four or five years they no longer have the problem. And in fact, one could say it's disappeared and then it will reappear at school age as a reading disability and therefore it is still significant. But a study from Whitehurst and your University at Stony Brook has in fact shown that it is a deficit that is predominately an expressive deficit. It probably is not significant.

The other comment I was going to make about the referral of the children with language disorders. The law requires now that all children who are at risk for developmental disorders be referred for intervention, so that the fact that so many were referred for intervention may be because people are following the Federal Guidelines which says between zero and three, if you have suspicion of a developmental problem, and particularly at this age language disorder is the first one that comes to attention, you will be referred for intervention. So I think that may be somewhat of an artifact.

Dr. Weil:

I work in the Early Intervention Program and I wish you were right, but in a study that we have done in Michigan, we think that there is less than 40%, probably less than 30%, of the kids who are eligible in terms of delay that are in fact referred for evaluation. Even then we don't know how many of those are getting treated.

The total treatment group for the under three in Michigan is currently 2% of the population, and that is probably up near the national average.

Dr. Davis:

I have a question for Phil. This has to do with his decision to look at stratification on clinics one by one. I am wondering whether you think it could be, in fact, done where you in essence look at the clinics by quartiles of some measure, or quintiles even. I guess I share the concern that you may have actually lost a lot of the informative risk sets by stratifying so finely on clinics.

Dr. Rhodes:

I think it's not so much that the risk sets have been totally lost, but that they are looked at quite differently in the sense that a clinic where 90% of the kids are in the same risk group. That's a higher risk group. The case being from that high risk group is certainly not an unusual event. Now you average that across a bunch of other clinics that

are very different exposure levels, that kid may seem unusual. But if stratification is the right thing to do, then throwing them back into a melting pot isn't really the right thing to do.

It's not good to increase your variance, but if everybody is the same in the group the kid comes from and you think that they are very different from somebody else, then they really shouldn't be there.

Actually I will say though, by the time I have included other kids that were excluded from another analysis and done everything else that I have done, the variances are really no bigger. Well, they are slightly bigger than before. The clinical stratification does increase them quite a bit, but doing everything else brings them back down to almost where they were, so it's not that I've just doubled the variances and that's why there is no affect.

Dr. Johnson:

Please, Dr. Clarkson.

Dr. Clarkson:

On another topic, you heard us toxicologists talking about body weights and what sort of blood levels we might expect in this population. Do the investigators have a Histogram of birth weights in this study? I haven't seen it. It might give us an idea of the sort of maximum blood levels we might expect to see in this group. To see whether these levels might overlap the lower ranges from other epidemiological studies.

Dr. Verstraeten:

All I can say about the birth weight, at least for the group of which we had the birth weight, was that the mean was 3.5 kilos and it ranged from one to about five, but I could produce you a histogram by tomorrow if you would like that.

Dr. Clarkson: Well, as long as you have a rough idea of what the range is, it would help.

Dr. Verstraeten: I can show you that by tomorrow.

Dr. Oakes: Everybody I'm sure is very well aware of this, but I don't think it has been explicitly mentioned. There must be a whole range of other potential confounding factors that we don't have data on. Can't measure predisposing to these various conditions. I guess it would be helpful at some point to kind of prioritize a list of what these might be and whether there is any hope of getting any kind of handle on it.

Dr. Johnson: Do you want to start the list of things that would worry you most?

Dr. Oakes: Well, I guess we have some data. What have we considered so far? We have some data on socioeconomic status. We don't have any kind of data on smoking, although it was mentioned. Alcohol.

Dr. Rodewald: In terms of what Phil Rhodes was talking about in terms of whether the low exposure group should be analyzed or not, that's a potential area to look for confounders. One of the characteristics that Dr. Modlin had mentioned earlier and the thing that's been troubling me the most, and that is that the lower exposure group are by definition late starters. There is a lot of health service research talking about the characteristics of babies who are later starters for vaccination and delayed vaccination. Many of these are socioeconomic factors and poverty access to care, which would not be a problem in this data set. But then there are subtle ones. This month in AJPH in basically the same group of children, there is a study shown that the late starters have less continuity of care. They see the same doctor fewer times after time, and that may have some

relation to being diagnosed with something subtle. So I think I would go to the Health Services Research and look for distinctions between things associated with late start that also may be associated with receiving one of these diagnoses.

Dr. Stehr-Green:

Let me preface this by saying this may make no sense because of my ignorance of the etiology and pathogenesis of development delays and so forth, but is there any merit in doing some sort of time series analysis to see if we can demonstrate a standard period between the point of exposure to the vaccine and the onset of whatever the outcome is? Is there any sense in doing that? The only temporal association you have demonstrated is that the diagnosis occurred after they first were vaccinated. Okay, that's the basic bottom line you have to demonstrate, but is there any merit in trying to establish if there was a unique or specific period of time post-exposure that development delays were first noticed? Does that make sense?

Dr. Rhodes:

I think Tom and I both have done some work on that in the sense that is there a different relative risk at age two as compared to age four, or one exposure group versus the others? I have not seen any big differences along those lines. I have not looked extensively or not across all the outcomes, but in the few I have looked at I have not seen anything that relates to that along those lines. I have also looked in different calendar periods. One thing that didn't really come out in the earlier discussion was that the diagnosis rates for these conditions, if you look at kids who are the same age at different calendar times, they have gone way up in the last five or six years in both Group Health and NCK. The diagnosis is much more common now than it was four or five years ago, for speech delays as well as ADHD.

Dr. Johnson: The comment was made this morning, or the question was raised are you just shifting the diagnosis to an earlier age and that if you looked later in time, would that even out?

Dr. Rhodes: I think the amount of risk that is being ascribed to Thimerosal exposure is so swamped by the calendar time factors. In other words, if there was an effective from Thimerosal and you stopped giving any Thimerosal, would you see a decrease in these problems? No, you would probably still see an increase because the temporal trends are so strong. Unless they have platformed out, they would swamp any effect of the size that has been contemplated here so far.

Dr. Johnson: But what you just said sounds like a conclusion. That the temporal trend is what is being measured.

Dr. Rhodes: No, no. I am saying that if one tried to do an ecologic study and said let's look at what the rates have been over the last few years. Now let's start not using any Thimerosal. Shouldn't they go down? No, they wouldn't necessarily go down because there has been a very strong upward trend in these two HMOs over the last three or four years in using, for example, the speech code, and that temporal trend may just keep going up and its slope is so much greater than the contemplated effects of the Thimerosal that if there were effects and you took them away, the trends may swamp that out totally.

Dr. DeStefano: I would like to make a comment on that. I am not sure if this was one the slides you had, but at one time I know you did some logistic regression looking at kids of certain ages. I think even up to six years of age. There was still a follow up, and then just looked at the proportion that had these outcomes. In that case, time of onset doesn't matter that much. You are just looking to see if they got it by that age. By the age of six, which for many of these conditions

including speech delay should be noticed by then. Is that one of the slides you had, the end results?

Dr. Verstraeten:

One of the concerns I had was the fit of the proportion of the hazard model. I am sort of surprised none of the statisticians have brought that up. I had a hard time trying to see if the fit was proper or not because the classical matters with all these strata was a bit hard, as I had more than 100 strata and I didn't really feel like doing it for every different strata.

This is by calendar year. This is still the proportional hazard model by calendar year, just to see if by a year of birth of the children to see if it was different for certain groups of children or not. So not all these estimates are identical. In general they tend to be similar. Note that in the last years the numbers become very small, however, in the first years the differences are not very large.

But if I can have the next slide, here instead of a proportional hazard model, we did a logistic regression model. I didn't use person time here and it's a bit tough to define exactly the control group. However, if I do it for all ages and not looking at different years, and this is for speech, the outcome is almost identical to the proportional hazard model, which suggests to me that it is not a question of bringing the diagnosis forward, but it is really the overall number that drives this estimate. And if I do it by years of the children, there is also hardly any difference, except above four years and then it sort of goes down. But until four years the estimates are not very different.

Dr. Oakes:

Those are cumulative, right?

Dr. Verstraeten:

They are not cumulative.

Dr. Oakes: So where they say one to two years, that is between one and two years?

Dr. Verstraeten: Absolutely.

Dr. Stehr-Green: But this is the age of onset, right?

Dr. Verstraeten: Age of onset or age of disappearing out of the group. The problem here is what is the control group I am going to use? So as a control group I used children that disenrolled, that reached the stop date before one year or between one and two years, et cetera.

Dr. Stein: Does this mean that speech delays were diagnosed under a year of age?

Dr. Verstraeten: Actually they were, yes. For some children they were.

Dr. Stehr-Green: That would be a very important point with regard to the accuracy of the study. Do you know how many?

Dr. Verstraeten: No, I haven't looked at that. I have no idea.

Dr. Davis: You and I both know it should be about zero, but it's about eight and it's children who aren't making sounds yet. Frequently these children haven't made sounds and they had an older sibling with a profound speech defect. I actually saw a couple of these and the parents wanted to make sure that they were sort of lining up the services that were available. At least at Group Health. That wasn't actually a rare scenario.

Dr. Johnson: If the number was small, hearing loss would be another possible explanation.

Dr. Rapin: I would like to make a comment. We have been focusing on all these acquired causes including mercury and

prematurity, and you had a list of confounding variables that should be considered in future studies. What we know today about all of the developmental disorders is that environmental factors are in fact rather unimportant in the case of these deficits and the major cause is genetic. So I think in future studies it would be extremely important that some genetic data be obtained. Questions such as is there anybody in the family who has reading difficulty? Because we know that the outcome of severe language disorder in the preschool child, the vast majority of the children will learn to speak, but they will reappear in the second bump of the condition as poor reading and spelling in childhood and adult life. I find it a little difficult knowing this and putting in autism. The major cause is not environmental, it is genetic and that we are focusing just on these environmental events or adventitious events when we haven't considered, and you told us that you don't have data for example on siblings, your study does not lend itself to considering the major variable.

Dr. Johnson: Well, I think the assumption is that those genetic predispositions would be randomly distributed.

Dr. Rapin: But you don't know that.

Dr. Johnson: No, that's an interlining assumption.

Dr. Rapin: I understand that, but you don't know that.

Dr. Johnson: Just on principle, Dr. Rapin, it seems to me that the more we learn about genetics or the more we learn about let's say autism, the more we shift toward focusing on genetic causes, but would you rule out the possibility, and let's move away from autism, that some of these are genetic predisposition and then the second hit?

Dr. Rapin: Not at all. I think it is in fact an attractive hypothesis.

- Dr. Johnson: Right, thank you. Yes, Bill.
- Dr. Phillips: I wanted to return to what I think was a thread that was just beginning. We talked about temporal trends and now we have talked about non-environmental causes. What is the population attributable risk we are talking about? Even if we assume that all children completed the complete series of immunizations and they all include all Thimerosal containing vaccinations, what is the burden of illness that we are talking about for these areas of interest? Speech delay and ADHD, that could possibly be attributable, if we believe these figures, to this exposure? What is the public health impact of the findings?
- Dr. Verstraeten: I haven't come around to calculating the attributable risk. I think it would be a bit tricky because we have different exposure categories, but I think it would be possible for each category to assign an attributable risk. As you are aware, however, a large majority of children are vaccinated, so it will probably be quite high, if we believe the signal.
- Dr. Oakes: On that calculation though, whether you choose zero as the baseline or the lowest or the largest exposure grouping would be a critical choice.
- Dr. Modlin: Two things. One I was just about to make the comment that I hadn't heard anybody use the term attributable risk for other reasons.
- Secondly, as a non-statistician, let me ask a naïve question. That is here we would have to assume if you use the term attributable risk, in part because the relative risk albeit may be significantly different are still extremely low, the risk ratios are low, that the true attributable risk is going to be low. It's just that when you apply even a very low attributable risk to a very large population, a large

denominator, then the actual absolute numbers become very important. Is that right?

Dr. Oakes: If you express it as a proportion of cases it is. If you express it as an absolute rate it would be, but as a proportion of cases which is fairly rare anyway...

Dr. Phillips: How about expressing it as the number of people per 100,000 population? My question is what is the public health impact of these findings?

Dr. Johnson: Could be large.

Dr. Verstraeten: Maybe to make a general remark on this, I have been a bit reluctant to get into such types of calculations. I think in the first place the whole face of this study was just to produce a signal, and what you are asking now is to extrapolate this to a public health level, which I have always been reluctant to do. I think in the first place that is giving credit it is not due, and in the second place, it is giving more accuracy to this data than what they really have.

Dr. Johnson: Dr. Snider?

Dr. Snider: I have two questions, but I will pose them one at a time. I am wondering from our mercury experts what they thought about Phil's presentation with regard to using I guess the 37.5 micrograms as a base, and then comparing the 50 and 62.5, 75 to that, and whether those differences in dose in the vaccines, whether based on knowledge of the effects of mercury, they would have expected to see the kinds of data we saw or something less or something greater? Given how much that dose would contribute to blood levels and tissue levels?

Dr. Clarkson: We went through this calculation last August, and that's why I am asking about the body weights. It would help a lot to keep us out mischief tonight if we had a reasonable idea. We have already heard what the range of body weights might be from one to five kilograms. It would be awfully nice to know what they would be from two months to four months to six months to get some feel for where these blood levels might lie.

Again, in doing such types of calculations, we have to assume it bears like methylmercury, which probably is not quite correct. But we could come up with some blood levels that would sort of relate to these that have gone on before.

It might be that in the very low birth weight group at the end of six months, we might start to approach some of the lower limits, where you would expect to see a small risk. But I don't know beyond that.

Dr. Johnson: I'm sorry, Bob, what was that comment?

Dr. Brent: I said the problem is the greatest risk for neurodevelopmental problems in a premature is the fact that they are premature, not the fact that they have gotten a vaccine with mercury in it. I mean it is very hard to sort that out. You know the high risk of neurodevelopmental problems in a 23, 24 or 27 week old premature.

Dr. Snider: I think I got the answer to my question. Basically that without knowing in greater detail the weights and being able to calculate the dose based on body weight, it is hard to know whether this is what you would have predicted or not.

The other question is for Phil or anyone who has had the opportunity to look at his analyses. That is if we turned the

problem around and said well, we have these data that suggest there is an association between exposure to vaccines and presumably the mercury component and these health outcomes, why is it that it goes away with the analyses that Phil has done? I am just trying to get a clear idea of the answer that question, because I think that is also important to be thinking about it from that perspective as well.

It is just hard to absorb all of this data at one time. I have had the luxury of seeing some of it more than once and I am sure some of the people who have never seen it are feeling swamped. The one thing I recall is the issue of the coding and some of the clusters of some strange coding, but I am wondering if Phil or some others would do a critique of that and say what might have made the effect for the moment we will assume is real, disappear with those adjustments he did? Would you be willing to criticize yourself, Phil, or do you want the other people to?

Dr. Rhodes:

I think some of what I did is not directly comparable to what Tom did in the sense that I haven't computed slopes if you will in that restricted range. For example, it is possible that even though say with five groups, I started with comparison group in four groups. A couple are significant and that goes away. It is still conceivable that you would see a mildly significant trend in those five groups, even though none of the four comparison ones are anywhere near significant. So to be strictly comparable with what Tom has done, I would have to go back and compute a trend statistic if you will for those three or five groups or whatever I have in my analysis.

I think the criticism has been made that maybe stratifying is too severe at NCK. I think the clinic is a variable that can't be completely ignored. It is going to require some looking at.

I think the biggest criticism of both of our models at this point is that they are over stratified on month of birth. That we really aren't analyzing as much of the data as we originally thought we were analyzing, and that some thought this is a great idea to control and an inadvertent in fact in a number of different ways, whose effects are still not totally understood.

Dr. Snider:

I guess to push a little more. It seems to me in addressing the question posed earlier about what epidemiologists are concerned about, and we have probably beaten on it enough, but I think we are worried about some kind of confounding in which there is an association between receiving vaccines or at least receiving vaccines on time, and the attention parents and health care providers might pay to their children's speech patterns and other behaviors, and therefore there is a greater likelihood of case ascertainment in that group that is well vaccinated. And there may be just as much disease in those that are not as well vaccinated, but there is not as good case ascertainment. That's the major epidemiologic concern I guess that I have.

But there is also this big junk of patients that fall out of the analysis, by excluding those kids who had some kind of diagnoses at birth. If I understand your analysis, Phil, including them seems to wash out the effect considerably.

Dr. Rhodes:

I think it has different effects. It goes in different directions for different outcomes at different times. It doesn't uniformly tend to bring things down.

Dr. Snider:

And I am confused as to how that happens. What could be going on that creates that kind...

Dr. Rhodes:

Again, I don't think there is any uniform effect that brings everything down. I think there was some that were actually

greatly strengthened by including those additional cases, apart from the fact that there were just more cases. There were some that went up and some that went down.

One feature that this does interact with potentially was clinic at NCK. There was one slide that showed the proportion of kids. The sequence of events here is that there are hospitals, birthing facilities at NCK, and they typically feed to one, two or mostly three clinics. So when you look at the birth facilities there was a huge variation in the proportion of kids that were excluded by these perinatal exclusions. That also varied by time for some reason. So putting these kids back in was a different thing for different clinics. For some clinics you were putting back in about 13% of their kids, or maybe about 10% because some would still be excluded, but for others it was putting back a quarter of their kids. So there was certainly a lot that could be going on in that sense.

Dr. Snider:

And I believe you pointed out that some of them were what we might term clinically at least, relatively trivial diagnoses and others were quite substantive diagnoses. They are very heterogeneous in terms of their clinical impact, and there was a broad range of the frequency in which various clinics ascertained these abnormalities. Therefore, as you point out, the percent that are withdrawn for those reasons varies considerably from clinic to clinic.

Dr. Rhodes:

Most of those perinatal codes are from the hospital discharge record from the birth. Now that is not always true. Some of these came from later ages and it probably wasn't even appropriate to their use, because they came when the kid was like a year old or something. But on the other hand, one hospital leads to two different clinics and there may be some relationship there between what the hospital will pick up and what the clinics will pick up.

Dr. Chen: To come back to the first point Dixie raised in terms of how do we get around this problem? A kind of health care seeking behavior bias potentially from parents. One of the reason that led me personally to not be so quick to dismiss the findings was that on his own, Tom independently picked three different outcomes that he did not think could be associated with mercury and three out of three had a different pattern across the different exposure levels as compared to the ones that again on a priority basis, we picked as biologically plausible to be due to mercury exposure.

Now Harry Guess kind of challenged us earlier to say that maybe those three aren't good outcomes. In which case then perhaps it would be useful for us to come up with what other additional ones do we want to test? If maybe five out of five or ten out of ten, all of those have a different pattern, then maybe that would be a way in which, based on these results before we do all these necessary studies, but those are going to take a much longer time, may be a quicker way to get at our answer.

Dr. Brent: Which one of the three that would not be associated with mercury?

Dr. Verstraeten: One was conjunctivitis, diarrhea and injury.

Dr. Rapin: Flat feet.

Dr. Verstraeten: There were two additional I added later that I thought would be more susceptible to parental worry, and that was flat feet and the code called worried well. Diagnosis not confirmed by physician, that is what it means.

Dr. Modlin: I know the hour is getting late, but one of the things we have not discussed is the date regarding premature infants in any detail that I found interesting. I understand that the

premature babies were analyzed. That these were babies that were just premature, but had no other diagnostic code. Is that the case? They were all the prematures, so they could have been associated with all sorts of confounding...well maybe that by itself is probably reason not to try to delve any further. But it was interesting that if you exclude the kids who got no vaccines for reasons that I think we all agree were likely to have been the most severely affected kids and therefore not immunized, it would be of some interest to note the timing of the vaccines for the other kids. Did the others get theirs on time or were they delayed? Or you would guess that they might be delayed. And whether or not you can't pull out of those groups the infants who had a diagnosis of just prematurity, but no other diagnosis such as developmental delay. Well, I am getting myself into a circular argument here.

Dr. Verstraeten:

It might be more comparable, what I have here is all premature, no matter whether they had anything else or not. Except they had to have two polio vaccines still. That was still there. So basically for like the entire category, the trend is even downwards. I am not sure what would happen if I took out the zero category. I am not sure if that becomes really flat or if there is still some kind of trend. However, the part that worried me was this. If I do the DTP-HIB combination, I can come up with relative risks of more than four, which is really very high. But the issue here is that as you can see at these confidence intervals, the numbers are quite small. We are still talking about 300 children. That is for the six months. The other one would be a smaller number.

Dr. Modlin:

It's unspecified delay, not a speech delay.

Dr. Verstraeten:

For the speech delay that doesn't happen. That's different. With speech delay it doesn't happen, so once again that is inconsistent here. I don't know about the prematures. I

think it is a very hard group to look at. If I take out the ones below 1500 grams, these risks come closer to one, but it is still not one. So that explains part of the effect, but not entirely.

The bottom line to me is you can look at this data and turn it around and look at this, and add this stratum, I can come up with risks very high. I can come up with very low risks, depending on how you turn everything around. You can make it go away for some and then it comes back for others.

To me the bottom is well, there is some things that just will never go away. If you make it go away here, it will pop up again there. So the bottom line is okay, our signal will simply not just go away.

Dr. Modlin:

I guess I'm thinking out loud here, but it might be that for some future study, that actually focusing on premature babies may make some sense, because we have all said from a biological risk standpoint, we would expect them to be at highest risk from exposure to Thimerosal if it was so, and it might be that designed to stay focused on those infants that don't otherwise have an obvious explanation for a cognitive problem might be a reasonable thing to do.

Dr. Verstraeten:

Maybe one thing to do would be to take the approach that Phil has taken. Saying that if they have reached 37.5 by three months, they probably didn't have a lot of problems, otherwise they wouldn't have received those vaccines. Although I am not sure because I can see some very premature children also getting vaccinated. I don't know, Frank, if you want to comment on other studies they have tried to do with prematures. It is usually not very straightforward. I was thinking of the neonatal mortality study. That was pretty impossible.

Dr. Johnson: Is it related, Tom?

Dr. Sinks: It's related to this. I caught the issue. The biggest concern we had with analysis was when they showed us the chart. If I am not mistaken that first no-dose category. Everything else was way below it and isn't that driving these numbers here?

Dr. Verstraeten: Yes, that's what I was mentioning. I think so, but I would have to look at those numbers, but anyway there is no upward trend. Of that I am sure.

Dr. Oakes: Well, I don't know that you can say that in looking at what those other figures look like. You can't say what it would look like without taking the zero group away. It might go up.

Dr. Verstraeten: I know it doesn't. I know it didn't. I'm not sure if it's above or below one. It's no different from one, but I'm not sure.

Dr. Johnson: Dixie, did you ask your second question?

Dr. Snider: Yes, I think I asked my second question, but I think I would just like to respond on the other diagnoses.

I think this was a reasonable effort and I, too, like Tom want to take the opportunity to congratulate the people who have been analyzing this data set. It's tremendously difficult working with administrative data sets and trying to make some sense of them. But I think it still might be worthwhile trying to give some consideration to some other diagnoses that might not have the same hard endpoints that conjunctivitis and some of the others did. I think flat feet and the worried well are reasonable things to look at, but maybe giving some additional thought to diagnoses that parents would consider as potentially serious problems, but

would not have the hard objective endpoints would make me feel a little bit better or worse, depending on how they came out.

I still remember when I first came into the TB research branch and had the opportunity with George Comstock to look a TB rates among people who had participated in BCG vaccine trials, but had refused vaccination with either vaccine or placebo. And to look at the TB rates, which should have been the same as the placebo recipients got, but in two large trials were about 50% different. And no one has ever been able to explain that. I think people who do clinical trials are aware of those kind of quirky things. I have realized that people who exhibit certain behavioral characteristics, whether it is refusing to participate or maybe seeking care more than other people, can have different outcomes and there can be disassociations, even though I guess we don't understand in this case the psychological mechanisms of the psychobiological mechanisms that are operating.

So it is not that I by any means want to dismiss this signal. As someone was talking about what is the attributable risks, there are tremendous policy implications for this. Not only as the issue was brought up with compensation, and we haven't heard from John Clements, but for global immunization efforts and so forth. But I think we have to be very, very careful that we got it right when we decide to make a policy call on this.

Dr. Johnson:

Thank you for that reminder. Yes, Michael?

Dr. Gerber:

Coming back to the methylmercurialization factor as a possible confounder, it seems to me that with the opportunity in this Harvard Pilgrim study, I don't know how much you want to get into this, but if there is an opportunity in that study to use as a reference group those

who had no Thimerosal exposure, not because their parents didn't commit for immunizations or brought them in late, but those who were not exposed because they received Thimerosal free vaccines. That would be a good way to try to deal with that. Is there that opportunity?

- Dr. Verstraeten: No, they pretty much used the same vaccines unfortunately.
- Dr. Chen: Five years after the VSD, we will be able to answer that question.
- Dr. Johnson: Could I ask, Dr. Rhodes said he was not too excited about trend analysis as it was used and I think Tom, you had commented a little on trend analysis, but how would you respond to that? Or Bob, either one.
- Dr. Verstraeten: Maybe I would have to ask Phil to clarify because I am not sure what his critique was on the trend analysis.
- Dr. Rhodes: I think my basic problem with it is that the assumption that 12.5 or whatever value you picked, 0 to 25, is the same as 25 to 50, 50 to 75 has the same effect. Unless you allow those separate groups to have their estimates first and you see they kind of fit a pattern that kind of adds up on a certain scale, to me going directly to that kind of modeling you can certainly obscure a lot of points. I have certainly seen cases where you have lots of ups and down, but you think you have a significant trend.
- Dr. DeStefano: Can I comment on how this unfolds? Basically as unfolded, the first presentation were just by category, and I think it around the second or third group that Tom presented when they started asking or started eyeballing. Saying this looks like a trend. Have you done trend tests for this? So then Tom started putting in test for trend and I guess with these big tables, it ended up being a convenient way to summarize data on those many tables. But it did

start out just looking at them categorical and then basically the audience kept requesting trend tests.

Dr. Rhodes:

I think it is important though to realize that the category models where you are comparing to the zero category are very different than the slope you go.

Especially since they are presented on the same overhead, I think people have gotten a little confused. They see the arrow bars. The arrow bars are for the model that compares each of the exposure levels to zero, but then there is a trend statistic on the bottom, so I think people have gotten a little confused.

Dr. DeStefano:

Then I am going to come clean on that as well. As Tom said, at first he tried to use the biggest group as the reference group, and then there was a lot of arrow bars that didn't overlap one, and I thought it might be better if we have more standard arrows that overlap one if it was going to get disseminated. I thought the zero group looked like a more logical release. When people want to see a zero group, I have become more convinced in the last intervening months that there is something pretty weird about this zero group or that probably Tom was correct to begin with, but still that's how it unfolded.

Dr. Guess:

I did want to support Phil's point on the issue of trend test and concern about them, especially the linear trend test. I believe there was an article in the American Journal of Epidemiology a few years back expressing a concern about it. So there are a bunch of different ways I think the statisticians could provide advice on how to do that, but there is a legitimate concern and some counter examples that show one can get confusing results.

Dr. Oakes:

I would just like to respond that you could always kind of produce examples against any technique. I am not familiar

with this specific article, but I think certainly the test for trend is a reasonable way to look at these and screen them as a preliminary thing, saying is there anything there or not? It is also a separate issue from where you include the zero group in there.

The other is you can do a test for departures from trend as well as the test for linear trend. I doubt that there would be enough power here to really detect any departures from the patterns that you do see.

But I think the other way of doing it and looking at each group separately and putting arrow bars on each group separately, you do dilute the strength of the relationship if you do that, so it's a trade off with power to detect a relationship.

Dr. Johnson:

For the non-statisticians, so when the confidence intervals consistently overlap one, but the trend is statistically significant?

Dr. Oakes:

I know these are not continuous exposure, they are actually discreet, but if you imagine it was continuous and you split them up into finer and finer groups with smaller and smaller numbers of people in each group, then the confidence intervals for each group would become wider and wider, and by splitting them up into enough groups you would get them all to overlap one, even if there was a strong relationship.

Dr. Sinks:

The way I have always used the test for trend is that it is not just simply a test for trend, but it is a test of the slope of the predictive curve. And that what you are really hoping is that you can use that test for trend when that curve actually is fairly good at predicting the point estimates for the categorical comparison. If it doesn't predict it well, it usually suggests you shouldn't be using the test for trend.

That is kind of my rough way of looking at it. There is a statistical way of doing that and I think Phil mentioned that in his analysis.

Dr. Oakes: It depends on the correlation between what you really ought to be using and what you are using. That is what governs it.

Dr. Davis: In defense of Tom, I think also one thing. First I agree with David in that one alternative is simply just to go back to comparing each category to the reference which does dilute out any signal, and you can then structure the categories to increase that dilution. But also I think one common sense approach would be to look at the observed trend. Here the observed trends haven't in fact been linear. We are not taking curve a linear or biorhythmic trends and doing simply linear trend tests on them. So I actually think there is considerable evidence here to support the use of the trend test from what Tom has done.

Dr. Braun: The zero exposure group, it sounds like Phil and Tom really chose different analyses there and they have an important impact I think on the results. I think there was some evidence presented with those kind of controlled diagnoses, conjunctivitis and gastroenteritis, showing a little step there, and that group was different even on those curves, as well as health care utilization and vaccine coverage at one year. I am concerned about that group as a comparison group. I was wondering, when you have the data analyzed two different ways, you could always present them two different ways, but somehow I think one way may be preferable, and I have concerns about using that. I would like to hear if Phil or Tom have anything to add about that.

Dr. Rhodes: I think my perspective was again, to look at the data and saw what questions do these data best answer? Not try to

start from the point of view that you have to answer this question regardless of whether the data is appropriate for it or not. And also taking a conservative point of view, that if you could see differences at 25 microgram levels, then you should be very concerned so that you wouldn't have to argue about these other groups.

So again, the fact that I don't feel like I am finding anything very strong in the data doesn't lead me to conclusively say that nothing is going on, but that beyond a certain level there is not a lot going on. In other words, that these 25 microgram differences, there is not a lot going on. And that whatever is apparently going on at the start, most of that is explainable through some other mechanisms, such as some of the exclusion criteria, clinic practices and that.

So obviously like my approach I did it, but it wasn't designed to give a definitive answer when it was negative. When it was positive, then I didn't think we needed to argue about some of the other aspects.

Dr. Brent:

I would like to tackle this question for my education. With the result that you have a slope over the period of time in the six months with regard to the results, what explanations would you have for that finding? In other words, all the ones that you could think of, of why you got those results? I have some explanations, but this is not my area. I would like to hear yours first.

Dr. Verstraeten:

You mean for the increased...

Dr. Brent:

For the slope of the increased risks with time.

Dr. Verstraeten:

What time?

Dr. Brent: Well, over the six month period. I mean many of your curves showed the rise in the relative risk, is that not correct? Maybe time. I mean over a period of time, you give me the explanation of why over a period of time you got this increased risk.

Dr. Verstraeten: I'm sorry, I'm not sure I'm understanding why you say it is increased risk over a period of time. Do you mean the risk increased?

Dr. Brent: Wasn't it true that if you looked at the population that had 25 micrograms you had a certain risk and when you got to 75 micrograms you had a higher risk.

Dr. Verstraeten: Yes, absolutely, but these are all at the same time. Measured at the same age at least.

Dr. Brent: I understand that, but they are different exposures.

Dr. Verstraeten: Yes.

Dr. Brent: What is your explanation? What explanations would you give for that?

Dr. Verstraeten: Personally I have three hypotheses. My first hypothesis is it is parental bias. The children that are more likely to be vaccinated are more likely to be picked up and diagnosed.

Second hypothesis, I don't know. There is a bias that I have not yet recognized, and nobody has yet told me about it.

Third hypothesis. It's true, it's Thimerosal. Those are my hypotheses.

Dr. Brent: If it is true, which or what mechanisms would you explain the finding with?

Dr. Verstraeten: You are asking for biological plausibility?

Dr. Brent: Well, yes.

Dr. Verstraeten: When I saw this, and I went back through the literature, I was actually stunned by what I saw because I thought it is plausible.

First of all there is the Faeroe study, which I think people have dismissed too easily, and there is a new article in the same Journal that was presented here, the Journal of Pediatrics, where they have looked at PCB. They have looked at other contaminants in seafood and they have adjusted for that, and still mercury comes out. That is one point.

Another point is that in many of the studies with animals, it turned out there is quite a different result depending on the dose of mercury. Depending on the route of exposure and depending on the age at which the animals were exposed. Now I don't know how much you can extrapolate that from animals to humans, but that tells me that mercury at one month of age is not the same as mercury at three months, at 12 months, prenatal mercury, later mercury. There is a whole range of plausible outcomes from mercury.

On top of that, I think we cannot so easily compare the U.S. population to Faeroe or Seychelles populations. We have different mean levels of exposure. We are comparing high to high in the Seychelles, high to high in the Faeroe and low to low in the U.S., so I am not sure how easily you can transpose one finding to another one.

So basically to me that leaves all the options open, and that means I cannot exclude such a possible effect.

Dr. Brent:

I think that is very helpful. I would add a couple of things in there and that is that there are three reasons why you might have the findings that you reported. One is, and we don't have the data, that with the multiple exposures you get an increasing level, and we don't know whether that is true or not. Some of our colleagues here don't think that is true, but until we demonstrate it one way or the other, we don't know that.

The other thing is that each time you have an exposure there is a certain amount of irreversible damage, and that as you exposure the damage adds up. Not because of dose, but because of they are irreversible.

And the third thing is that maybe the most sensitive period is later, like in the fifth or sixth month. In other words, the sensitivity period is not the same over the first six months.

Those would be explanations that you could only demonstrate with research, and probably not human. One of the things that could be done here, since we don't have a lot of human populations and that is going to take a long time, is to model an animal experiment.

I was involved in the allegation that came from the ABCC that one rad of radiation resulted in a doubling of the incidence of mental retardation, which didn't make any sense. We went back to the laboratory and did an animal exposure using nine neurodevelopmental behaviors and showed that at one rad, you have no pathological effects. The central nervous system was effective and the neural behavior was normal.

I think the government could put together a project like that, just to see what the threshold is for neurobehavioral effects. You can't use the rat to predict things in the human, but it could give us some information that would be

a little helpful. Because the big problem is all the things you say about mercury are true, except the fact is it is important on the dose and we don't know what the threshold is on mercury. If we are below the threshold for any effect; then all the things you say with regard to the toxicity of mercury are just not valid, but we don't know the threshold dose.

Dr. Johnson:

Bob, when you focus only a threshold, you make the assumption, isn't that kind of a puristic constant. You make the assumption. When it reaches a detectable point across a population so that when you are dealing with human beings you have a lot of different genetic make ups and presumably you get the end large enough and those are blanked down, but if you look at individual cells you add these things and they affect the cells. Each individual cell. So you are focusing a lot on what you can measure as your endpoint and your determination that yes, this is a threshold effect. There is no gradient effect, and that worries me. In general it worries me. In any kind of assumption that this is only a threshold.

Dr. Davis:

But I think just to take what you were saying a little bit further, one could posit that there is a normal distribution of background mercury in the human population, and by vaccinating everybody at one single time you have raised that and in essence moved that entire normal distribution some segment to the right, and you may in fact get some very small, but detectable portion of that population in the affected range.

Dr. Brent:

Well, I don't know.

Dr. Snider:

Just to build on what Bob said, based on earlier conversations about this population. One might support a hypothesis further by saying that if the people who are more likely to be on time for their vaccines are the higher

socioeconomic group and they are like the health care workers that Dr. Koller showed us and so forth, it may be that their baseline levels of mercury are higher. So what you are doing is seeing an effect in a population that has higher baseline levels of mercury, and you don't see it necessarily in those that are lower because first of all they are not getting the extra mercury anyway. So you have exacerbated the problem even further than what you have just described.

Dr. Brent:

The implications from this discussion is that the threshold is very near what we are talking about here. The fact is that we don't know that. You might find that it is tenfold or even a hundred fold higher with regard to some of the things we are discussing. You have to do the study, whether it is in the human or in the animal. I mean all these hypotheses are valid hypotheses to test, but I can tell you in our field we don't have a single agent that produces birth defects of the central nervous system or any other organ that doesn't have a threshold. If you want to make birth defects with Thalidomide, you give 50 milligrams. You can give every mother in the world one milligram and nothing will happen, and that is true of every known teratogen. That is a typical toxic curve. That is because it is a multi-cellular phenomenon. It is a toxic phenomenon. It is not a stochastic phenomenon. We need to data to answer the questions that you are raising.

I would ask our mercury experts with regard to the fact that I don't think there is a spectrum of genetic susceptibility to mercury like there is to Dilantin and many other drugs with a bimodal curve. I think that is a very narrow spectrum. I would like to hear from our two experts.

Dr. Johnson:

We are going to have to close and I will let Dr. Sinks have the last comment, but Bob, let me just try to explain a little further my concerns. If you look at fetal alcohol syndrome

that started out as a very striking syndrome. Facial and so forth, the better we have gotten at analysis and broadened our analyses, we decreased the fetal alcohol effects and further to possibly ADHD and so forth. So declaring that you have a threshold effect, it continues to worry me. You and I can discuss it later.

Dr. Brent: Fetal alcohol effect. This is 50 milligrams per day. When it gets a little better, you don't have any effects and that's only a glass of wine a day.

Dr. Johnson: I want to see what the endpoints are. Tom.

Dr. Sinks: I was going to say something similar. I think it is fine when you are comparing apples to apples and you are saying the effect of this. This birth defect, phocomelia for Thalidomide, but when you are looking at something and you are changing your effect and you are looking for more subtle effects, and lead is a classic example where we are looking at more and more subtle neurologic outcomes, we start dropping down what that threshold might be. Because we have changed the way we are measuring the outcome. And as long as we have faith that the outcome we are measuring is real, then that threshold is changing on the basis of what the outcome is we are measuring. So there are two things going along at the same time. One is the outcome, and the other is the threshold. You are kind of keeping that threshold as a constant based on the outcome.

Dr. Brent: I think they have been saying that there is a threshold, and I would like to know what it is. I'd like to probe and find out what it is.

Dr. Johnson: The smarter we get, the lower the threshold.

Dr. Bernier is going to allow us to end and go out finally into the fresh air.

Dr. Bernier:

Quickly if I may, I want to talk about the homework assignment for the consultants, and I would like to invite the other members of the meeting to feel free to fill it out. We would like to collect your opinions, although the people we are obviously looking to are the 11 consultants. I am sure if you have seen the list of participants, you know who you are.

I just want to read these questions in case there are any semantic issues, because we did focus carefully on this and we don't want to have any semantic problems when the questions are answered, and then oh, that's not what I meant. So let's try to make sure that we understand clearly the questions.

There are three questions altogether. As I mentioned this morning, I would like to suggest that you take your notes this evening and make notes on here as preliminary answers. Use that tomorrow morning to make your comments because we will go around the room person by person. There are 11 consultants whose opinions will be solicited, and then after you hear those opinions, you may want to make some revisions on the final sheet you turn in.

DAY TWO

Dr. Johnson:

Are there any questions that anyone wants to pose the presenters from yesterday?

Dr. Braun:

I have a question that did not get answered, but I don't see Dr. Verstraeten here. It had to do with the presentation of data and Dr. Rhodes was concerned particularly about using that first zero group as the reference group, and all the relative risks that were presented on the graphs were based on a comparison to that group. I think that might be a pertinent issue to the extent that the data are presented here, but outside, and will affect the risk estimates. I

thought that might be a useful thing to consider, but I don't see him here, so I don't know.

Dr. Chen: Well, he should be on his way, so why don't we go on to other questions.

Dr. Clarkson: You mean we are lost? People can't find us?

Dr. Davis: The only thing I do know is that I think on page four of his hand out that he had Thimerosal and then showed analyses. It has sort of a cryptic title, but I think when you kick out the zero exposure group, the relative risk to language, speech and unspecified delays seems to remain relatively unchanged. It's kind of hard since we had so many analyses we were talking about yesterday.

Dr. Braun: The thing is, if you present the data as trends, but if you present the data with the arrow bars and the real risk, then I don't think, so it depend on how you want to present the data. And then if you do present then with each stratum, each category having its on relative risk, then it would affect the risk estimates. But not if you present the data as just a trend with one number characterizing the trends.

I guess I shared the concerns that Phil raised. I thought they were valid and convincing. He left more leeway with talking about the next two or three exposure groups and said there may be some value in those. But the zero group seemed to be different and many of the analyses, even the ones where it shouldn't have differed, so in my opinion if you are going to do those categories versus one reference group and then every category you look it in comparison to that group, in my opinion those are not useful to present.

Certainly there is a high risk that they are biased, so I just wouldn't recommend those. I would be interested in other people's opinions. So rather than recommend a specific

way to do it, I would you could either do a relative risk of the strata versus a difference reference group. Or like Bob was saying present a trend number which from what you are saying, that wouldn't change if you want to do it that way. I don't have the same problem with doing it that way. I hope that is clear.

Dr. Caserta: I have a question for Bob. When you did the chart review, Bob, did you look at the zero group to see if there was any obvious difference with that group as opposed to the rest of the cohort, or was that not done? Can you describe the zero group in any other way other than saying that there is a zero group and that's all I know about them, and that they had two polios?

Dr. Davis: No, we did the chart review completely separate from exposure. We literally had no idea what the exposure was on purpose and I provided the chart review. Your point is well taken.

Dr. Clarkson: Mr. Chairman, when I look at the paper here, the graphs don't always say zero. The reference. They say less than 37.5, then say less than 25, so are they all referred to zero or not?

Dr. Braun: No, but even in those less than 35, they are part of the less than 35 group. I mean they could be excluded from that.

Dr. Chen: I think the bottom line is that while the zero group is different, and I think all of us would agree with that, the issue is that it is impossible, unethical to leave kids unimmunized, so you will never, ever resolve that issue. So then we have to refer back from that.

If we can never, ever leave kids unimmunized through these age groups in order to study them ideally as we would like, then of the kids who did become or are left

unimmunized for whatever reason, be it that their parents are socially responsible or be it that they have some other pre-existing condition medically, we just have to work with that.

I think if we throw them out, or maybe I think the thing to do is that I would chart review, I guess it would make sense on the chart reviews to focus on those cases a bit more to understand. If these kids are otherwise normal and they really are just not being immunized because of social circumstances probably, we need to make some judgment as to are they otherwise at risk for the outcomes that we are looking at.

Dr. Caserta: But Bob, a study could be done. You could use the acellular pertussis that doesn't have Thimerosal in ComVax, and have children be immunized, but not have any Thimerosal.

Dr. Braun: Sure, we will have the answer in five years. The question is what can we do now with the data we have.

Dr. Chen: One of the things that Vito said was how were they different? I think there was a graph of the health care visits in the first year and they had fewer health care visits than the other.

Dr. Braun: But if that is purely because they come from parents who otherwise are busier or whatever reason, but the kid otherwise is normal, would you want to throw them out? Are there criteria we could develop when we go to the chart review that would permit us to retain them?

Dr. Stehr-Green: Well, I think the issue is whether or not they have the same opportunity. If they were to develop the interests, if they were given the same opportunity to inspect them, I think the answer is no. And I'm not sure any amount of chart

review is going to resolve that issue. They are fundamentally different. They have differential potential for ascertainment, and I don't think the chart review findings are going to resolve that.

Dr. Caserta:

But you don't know that. You don't know that there is differential potential. They may have gone to the doctor less because they weren't as abnormal. You just don't know.

Dr. Stehr-Green:

I don't know the reason, but the evidence was presented. There were several evidences that were presented that suggested that there was lesser opportunity for them to have been affected. They sought health care less frequently than the higher exposed groups, and maybe it was because they were healthier and they weren't affected by Thimerosal or whatever. But the fact of the matter still remains is that they did not have the same opportunity, if they developed these outcomes, to be affected.

Dr. Hadler:

I basically agree with the issue of how you handle the less than 37 group in this analysis you presented. I think it has to be thought through carefully. As you can see, a lot of the analyses, when there are fewer outcomes they have already lumped them together. In others they have kept them apart, and yet the numbers of outcomes in much, much smaller than it is once you get up to 37. The real issue, is the zero group very different, which it appears to be in some analyses. The 12.5 and 25 are less clear, but do you really have the power to discriminate between these three groups or is it better to always keep them lumped? It sort of gives you a false sense of well, we can say there is a linear trend beginning at zero and going up to 67. And yet you really just don't have the power, even if your biggest, for these lowest exposures and I think very careful thought with the statistician has to be given as to whether you keep

presenting it with those three split apart or whether you group them together.

I think someone also needs to look again at these groups as carefully as you can to just know as best you can how much they are different in terms of health seeking behavior. Part of it is probably spelled out in some of these tables, but get the best understanding you can to make some decision. Are you going to try to split them apart and give a full sense that there is a difference between them, or whether there is a power to differentiate or to see a difference between them, or perhaps just lump them and say we cannot say below 37.5 that there is any difference among those group. So for the purpose of this analysis, we are going to put them together.

Dr. Rodewald:

I think he may have just made the same point, but it's not just the zero group, but the first two are clearly late starters because of the first dose of HepB. And actually the first two and three quarters of the groups are really late starters, so it is the three groups. We have been just saying the zero group, but it is more than the zero group.

Mr. Schwartz:

The thing is we have all looked at the fancy epidemiological analyses. To me one of the most important pieces of data presented was the crude incidence rates. The outcomes which was on page 9 of the original material that Tom presented. It shows if you look at the incidence of speech delay and ADD, it shows that these outcomes in the zero group and the 12.5 group actually are diagnosed more frequently than in some of the groups that have higher exposure.

I think the other thing that really stands out is that if you look particularly at Group Health, there doesn't seem to be much of a trend toward those increased outcomes with increased exposure.

I talked with Tom during one of the breaks yesterday and what he mentioned was if you do the trend graphs for Group Health alone. In other words, if you separate Group Health and Northern Kaiser and do those trend graphs, that with Group Health you see an increase from the reference group. From the first category to the next and then straight lines. So you don't see a trend with the Group Health data separated from the Northern Kaiser data. And that really the graphs that were presented are driven by Kaiser, which has a much larger patient population. So I think one of the points that is worth making is that the information we are basing our conclusions on are really more related to a single managed care organization rather I think that the combination of the two. And if that is not correct, maybe Bob could indicate that, but I think that is correct.

Dr. Davis:

I am uncomfortable having to speak in Tom's absence, because he knows the data certainly better than I do. But I do know one problem is simply that they are crude. So I agree with what you are saying, in pointing out that they are crude. And as it pertains to the combined graphs that we saw yesterday, you are right. Whenever you combine a gorilla and a small mole, it is going to look mostly like the gorilla and that is what we are seeing. Northern Kaiser has always been bigger than Group Health and there are many other issues attached to that.

But nevertheless, when we combined the data it is almost...

Mr. Schwartz:

Then the Group Health data are essentially clad across the different exposure categories?

Dr. Davis:

Well, they have a different appearance that varies by disease and here is the man himself, but I don't think it is proper to think there is no trend at all. It has a step wise trend and then a flat.

Mr. Schwartz: But that first step is the same thing we saw, for example, the first step going in the opposite direction with the prematures and then it was flat.

Dr. Bernier: So what that initial study was, but I would hesitate to make any analysis of the prematures frankly because I think there is so many of those confounded by variations in the prematures.

Dr. Weil: I don't think there are sufficient data, not just in this study, but anywhere to make the assumption that ascertainment depends on the number of visits. It is sort of a reflex concept, but in fact having done a lot of work on trying to ascertain developmental delay with kids under three, we can't find that the number of visits determines the rate of ascertainment. And not in these data, but in other data that we were working with, so I think people jump to that idea because it is intuitive, but the fact there are no data to support that concept that I know of.

Dr. Bernier: Jose? Dr. Johnson had to step out for a moment for a personal call, so I will step in for him until he gets back.

Dr. Cordero: I'd like to follow up on Mr. Weil's comments, but if we look at page 5 of the additional analysis hand outs, one of the things that I was impressed with is that the group on the bar charts, and basically the groups that have 0, 12.5 and 25 micrograms all seem to have completion rates of 60% as compared to the groups that have 37.5 or greater. In the National Immunization Survey, when you look at the risks on their immunization, there is some 4-P risk factors. One is maternal age less than 19, some lower socioeconomic status. Meaning a family below poverty line. Number three is households with five people or more, and the maternal indication was in high school. That is linked to maternal age, too. All of those factors also tend to be

related to the fact that the parents are going to likely to be paying less attention to especially subtle abnormalities.

Often because the children are going to have a visit doesn't mean they are getting immunized, nor are they getting diagnosed. Especially things as subtle as some of these developmental disabilities that may not get picked up on a single visit.

Dr. Orenstein:

It seems to me that when you have such small percentages of the population getting zero, 12.5 and 25, I have a fundamental discomfort of trying to say that group is a very strong referent group to the rest.

It seems to me the strongest data begin at 37.5 micrograms. As Jose pointed out, that is the group that was finished on time, even though they were perhaps starting late. I think that if the trends are there, 37.5 is your reference group. Those I think are perhaps more concerned than if you have to start at zero. We all realize that zero is problematic. We saw it conjunctivitis and in others. It seems to me from the scientific perspective, 37.5 as the referent group makes sense.

Dr. Bernier:

Can I ask if some of the statisticians or epidemiologists if they want to comment on that, and then move on to ask the individual consultants their opinions, but David Oakes, do you want to comment on this issue?

Dr. Oakes:

I do find it a little confusing that the groups switch around. The reference groups switch around with the different diagnoses, and the reason is the end is different and there weren't enough people, but it does make it a little hard to compare across, so I would advise trying to make it consistent. But certainly I think we are agreed that the zero group is not a good group.

Again, I do want to emphasize that if you are doing the test for trend, essentially that does not use a reference group, so it is one argument for analyzing the data that way rather than computing relative risk.

Dr. Bernier:

Anyone else? Paul? Dr. Kurz?

Dr. Kurz:

I have trouble, too, with the referent group because by using this zero exposure, because there is a lot of difference between zero and 25 and the other exposure group, and when we use this zero exposure as a confidence variable, they are very much an influence. They also influence the P-value. I don't see a curve with a fitted variation nine with an exposure to see what was really the fitted line by using the zero exposure, because it may be an interest if you use all the diagnostic criteria to test this relationship and...

Dr. Verstraeten:

No, I haven't. One thing I have done was to take out the zero group and that does not affect the estimate. The side effect group is so small, it really is a very low influence. If you start taking out the lower groups, I know for speech I could take out up to 25, even I think 37.5 and the trend would still be there, so at least for that it doesn't affect. I have not tried it for all the other ones. So I am not sure what the effect would be for the other ones, but I would to reemphasize what data that has. Once you have the trend test, the influence of those category groups is quite small because they are quite small in sample size and they are not a reference group anymore. There is no such thing as a reference group at that moment.

Dr. Rodewald:

In my mind, are we talking about taking out the bottom three groups? The below 75? Is it as if you could put your thumb over those points and then take a look at the rest of the line and say that that's what it is, or just this really do a reanalysis of the different referent groups and then you

may have something that is no longer the same line? Because I think, at least in my mind, maybe not others minds, but it's just like if we put our thumb over those lines, problems just go away, but I'm not really statistical enough to know if that is true.

Dr. Oakes:

There is a difference between taking them out and combining them. I'm not quite sure which we are talking about. Are we talking about putting all these three groups together?

Dr. Braun:

I think you still have to be careful, even with the trend, because the trend line is saying for every increase in milligram of mercury, you are increasing the risk X, and if the data is really based at 37 to 75, then if you talk about the zero to 37 group you are kind of extrapolating. Okay? Because if the line is really coming from that range of data, then people are going to turn around and say there is 75 milligrams, they are really going to take into account the beginning of the line. So isn't that kind of an extrapolation? And even Alex Walker was saying yesterday if I heard him right, that those data are not influencing the line. The lower than 37 because of the small numbers. So in a sense there is some extrapolation that is going on. So in a way it is more satisfying to use the trend, but you are still not totally obviating the issue.

Dr. Ellenberg:

I think we all face an interesting point that you don't get into the people who actually got their full vaccination series until you get to 37.5. I guess I would worry a little bit that we started here and left the others who might be trading one bias for another. Because then we have a group of people who got their full vaccination series by the end, so why were some of them late starters? And might that relate to their medical status? I don't know why, whether it would be a random thing that some started late, so I don't

know what possible bias there would be with that approach either.

Dr. Weil:

I think there are two things we need to think about. One is, is there a threshold? Bob Brent raised the question yesterday. Are we talking about a threshold effect or are we talking about from zero, so the first microgram of mercury has an impact. It seems since we are not talking about malignancy that we might very well have some kind of threshold phenomenon, so those values that are down below that threshold may in fact be of very little consequence. We don't know that and we will never find that out from these data.

The second point is there is something else we won't ever find out from these data, I don't think, and that is whether 37.5 milligrams at one month is different than 37.5 milligrams at two months or three months, and that may be because of brain development. A critical issue and we can't answer that either from these data, no matter how they get manipulated or how many times we review. So some of the really gutsy questions from a person who is very concerned about neurodevelopment cannot be answered out of this. I don't think we have anything that says this establishes this. All we can say is we are anxious, and we need to get data the way we ordinarily do. We need to go to animal neurotox studies, developmental neurotox. We need to look at some other data that can be obtained to see if we get a comparable kind of impact, but let's not try and refine and refine and refine these data. These are what they are. They show something and you cannot, by twiddling them and manipulating them, get much more out than Tom, Bob and others have already done. They've done an amazing amount with relatively little data, and I think I am impressed at how much they have got and I don't think we are going to get anymore out of it.

Dr. Johnson:

Thank you, Bill. I think that is a good transition comment so that we could go on and move into statements from the consultants. We are going to go question one and we will then have question two. We will have a presentation before we go back and deal with ideas about research. Maybe I should read the question so you don't have to do it. Do you think the observations made to date in the Vaccine Safety Datalink Project about a potential relationship between vaccines which contain Thimerosal and some specific neurologic developmental disorders, speech delay, attention deficit, ADHD and developmental delays constitute a definite signal? That is are a sufficient concern to warrant further investigation? Paul?

Dr. Stehr-Green:

First I wanted to reiterate what others have said. I want to congratulate the folks who did the initial analyses for a tremendous amount of work, a lot of dedication and very interesting results.

In my judgment, these preliminary results are not compelling, but the implications are so profound that the lead should be examined further.

My outstanding concerns and reasons for that statement really go to the validity and the accuracy of these results that revolve primarily around the issue of ascertainment bias or confounding, which I think is potentially a fatal flaw which was not dispelled by some of the clever analyses.

Some other concerns I have deal with the uncertainties, as we talk about the low dose groups, and I think Dr. Rhodes demonstrated those concerns very nicely. In effect that is closely related to the first issue of ascertainment bias.

Another concern I have is the inconsistent and in effect mostly unknown case definitions that again, even though

Dr. Davis did a very nice job of going back and showing that at least for some of the major outcomes, that the initial information on the electronic records were very closely supported by more detailed clinical follow up, I think there is still a major issue of is a case of ADD a case of ADD everything, at least as ascertained, over this time period?

Then finally I think as Dr. Rhodes pointed out that the exclusion criteria may have introduced other biases that have altered our ability to draw inferences from this data.

Dr. Brent: One of the things, I think we might allow a couple minutes of discussion to clarify some of those points. I am not sure I understood. Are you voting yes or no?

Dr. Stehr-Green: Voting yes, the implications are so profound these should be examined further.

Dr. Brent: So the reason for voting yes was sort of a show of problems rather than the reasons we should pursue it? You gave limitations of the data rather than explaining why you think we should conduct further investigations. Unless you have one basis reason, which is not the data, but the implications of the data. Is that right?

Dr. Stehr-Green: Yes, and I guess what I wanted to talk about were those facets that...

Dr. Brent: The problem being with the data, is that right? They don't really explain why you think it should be further pursued. The main reason that you think it should be further pursued is?

Dr. Stehr-Green: The implications and if further research is done, I hope that it can somehow rest these concerns or mitigate these concerns.

Dr. Brent: So the reason for further investigation is not really from the data themselves? It's not in the strength of the data?

Dr. Stehr-Green: Not on the strength, no. They are intriguing, but certainly not compelling.

Dr. Johnson: Paul, some of what you said might fit under question two?

Dr. Stehr-Green: Yes, all these questions were kind of inter-related.

Dr. Johnson: Kevin?

Dr. Sullivan: I said yes. In my mind it appears there may be a small possibility of some increased risk. I am not convinced that there is, but the question was do we stop and not do anymore work, or should we go on and do some further investigating? I say that there should be some additional investigation into the potential association.

Dr. Johnson: That seems clear. Dr. Clarkson?

Dr. Clarkson: I said yes, too. I am not quite as enthusiastic. I only heard Dr. Weil's comments, but I was giving the same reason that maybe some additional observations could be made. For example, some of the non-mercury endpoints could be looked at. Again, I come from a long line of researchers. I hate to say no to stuff in research.

The point I think is unique from a mercury point of view in that there is an astronomical number of people in this study. All previous mercury research has involved epidemiologically groups of less than 1,000 infants. To go from 1,000 to 100,000 is a staggering jump. So I am fascinated by the size of it.

Now if you take out the Faeroes or the Seychelles, although they disagree as far as prenatal outcomes are concerned,

they are in agreement in terms of postnatal outcomes. All indicate that they cannot find any fact due to the postnatal exposure. For example in the Faeroes, they looked at the kids at 12 months of age and found that the higher the mercury levels in these kids, the more rapidly they obtained the developmental milestone. If you recall, they confounded or they suggested it was breast feeding. There is a lot of the breast feeding theory. The higher the mercury level in the kid, presumably breast milk being the source of the mercury, and of course the benefits of breast feeding. So what they found in the first 12 months was they could not find an adverse effect.

At six months and at 19 months in the Seychelles, we couldn't find anything either. And in Iraq where we looked at kids with astronomical blood levels, up to 1,000 parts per variant in blood, well experienced pediatricians as a team could not find anything obviously wrong with these kids. So the recurrent body of evidence says that postnatal period is not the window of susceptibility. As Dr. Brent mentioned yesterday, it is probably to do with neuromigration, which is in an earlier part of gestation.

On the other hand a point I have said before, that these studies are 1,000 or less and here we have 100,000 infants, so as a mercury manic, to make me say yes, let's keep looking at this group, it's a very large group.

A third reason for us to continue is that it might be a guide to future studies. I don't know whether future studies are possible, given that maybe mercury in vaccines is coming out now, and maybe not in this country, but elsewhere. It might guide us to what other additional things you could look for in a future study. For example, the role of breast feeding probably is very important in determining these outcomes. And of course you can't get it in this particular study, but maybe in a future prospective study you could

look at that. So these are my reasons and I expect to get 10% of the budget.

Dr. Johnson: By the way, my understanding, the current understanding is that neuromigration incurs even in adult brains, and that this has been shown in animals and it has changed the whole concept of spasticity.

Dr. Rapin: But we are not talking of the migration which results in the organization of the cortex, and the amount of migration is small and it is horizontal and not vertical. It is a completely different phenomenon.

Dr. Johnson: It is not the same thing, but I don't think it is correct to assume that there is not a whole lot going on in the central nervous system from the time of birth on.

Dr. Bernier: I mean compared to what went on in the embryo, I think it is miniscule. All the cells that make the neuron come from that single cell layer, the appendum of the brain. They are gone. They are not there anymore in the adult. You can't form any new cells from the neuroplast. So what you are talking about is an interesting phenomenon, but we don't know of its implications with regard to behavior or learning.

Dr. Johnson: There is a lot of study in that area, Bob. In any case, we don't need to get into this. David?

Dr. Oakes: With regard to the first thing Tom said, in 30 years I don't hear everything and any group of experts addressing any topic when the group has felt comfortable at the end of the meeting, saying we know everything we need to know about this. Let's not do anymore research. It would really go very much against the grain to take a no on one. And that is not related to the strength of the evidence. It hardly

matters. Actually the methodological issues and the interest in the topic.

Also I think things could be done, further analyses of the data, further confirmation of some things and not at great cost, that would help clarify at least some of the issues involved.

Dr. Bernier:

One of the reactions I am having as I am listening to this, I agree with you completely about this 30 years and never expecting scientists to say that they don't want to do more studies. That more studies would be good. So I am wondering why question this? We knew the answer, so let me try to defend the question a little and if you agree, maybe we could start over again.

The point I am making is that the way this question was written is not do we need to know more about mercury? The question is really do you think that the observations that have been made in this project are of sufficient concern that you want to investigate more the relationship between the vaccines which contain mercury and these outcomes? So it is not just a question of do we want in general? It has to do specifically with that issue. Is the level of concern that you have about it, has that been raised enough by what has been observed that you want to investigate more that specific question? I don't know if that is the same thing.

Dr. Orenstein:

I think, Roger, that is the same question. I think perhaps what is a better question, is what is your level of concern about these findings?

Dr. Bernier:

Well, that really is the second.

Dr. Orenstein:

Well, the second issue is we don't know causality. We don't know about causality, but is this something that really warrants some urgent attention? It is two issues as to what

is your level of concern and the need to look further in terms of concern, whereas I don't know how the people got causality. But there may be a different issue as to whether they are willing to admit to how strong plausible versus this is something we need to worry about and we need to do more on it.

Dr. Johnson: Roger, would you like for us to grade this as 1+ or 4+ so you don't get too fine a cut in terms of concern. Would that introduce...

Dr. Bernier: Let me stop for a minute because I am trying to think about the point that you are raising, Walter, and it seemed that it wasn't helpful to just hear about the level of concern because to interpret that, it could have multiple repercussions. It could mean that it is concern, therefore that concern needs to be translated into a policy action or it means that the concern is that you don't think the evidence is strong, and therefore it is not worth doing more research. I mean just to measure people's level of concern without trying to get a handle on what does that operationally mean, I don't think is really helpful. So the reason we put this question this way was to operationalize what was meant by the signal. And likewise by the second question, it was to operationalize it by expressing it in terms of what you thought about how much this supported a causative relationship or not.

I don't know if others have different views and I don't want to get into a big semantic debate, but on the other hand I don't want to wind up after the meeting and people feel well, we could have fine tuned what we were doing and it would have been a little more helpful. This is a rare opportunity we all have to be together this morning to hear one another on this, so we want at the end to feel that we got the most out of this. So Susan, do you want to make a suggestion?

Dr. Ellenberg: I may be jumping the gun, but one of the ways you could frame it is the level of concern sufficient to have an urgent and immediate research plan to address the question. And the other one is the level of concern sufficient to instigate a chain of policy? I know that's jumping. The best way of measuring the magnitude of concern as opposed to measuring it related to causation, which I don't think anybody would be able to say that they know.

Dr. Bernier: They don't have to know, they just have to render. The way the question was written is that you render an opinion about the evidence as it exists. Does it or does it not support a causal relation? It is not a yes or no question, it is just that how much do you feel it does support it?

Dr. Ellenberg: But I think in terms of quantitative concern would at least may be able to determine what kind of action you can take.

Dr. Orenstein: I think you are talking about two qualifications. One is what is the level of concern of the need for action? I agree, I think I would be shocked if everybody went around the room and said I'm just not sure. May be, but I think the issue is what is the level of concern.

Dr. Bernier: But what will that mean, Walter, if after this meeting everyone goes around and says I have a level of concern and it's high. What are you going to do?

Mr. Schwartz: Can I just make one quick suggestion? In the past you asked the question how strong are these data as a signal? That might be one question. What do you think of these data as a signal of a problem?

The second question might be what is your level of concern, and concern brings into account the signal, but it also brings into account all of the expertise that these mercury folks have given us, and these developmental

folks about concern regarding not just the signal, but the issue in general.

And the third question is what do we do about it and is more research needed and how urgent is that research? So if you are trying to separate what do you think of these data from what do you think of the issue, that might be one way to do so.

Dr. Oakes: The other side to this is these data are now out. I mean they may not be public, but they will be. So this data exists, and then we can't go back to the state where this duty has not been done, so there is a need to understand the data we have. And that might be the way I would frame it. A better understanding of the results that we have.

Dr. Clover: Maybe that is an impossible question to answer, your first question, because no one around here is going to say that mercury per say is not a concern.

Dr. Clarkson: Thank you.

Dr. Stein: Let's go on around on the first question.

Dr. Weil: I may have helped or not. My answer is yes. Although the data presents a number of uncertainties, there is adequate consistency, biological plausibility, a lack of relationship with phenomenon not expected to be related, and a potential causal role that is as good as any other hypothesized etiology of explanation of the noted associations. In addition, the possibility that the associations could be causal has major significance for public and professional acceptance of Thimerosal containing vaccines. I think that is a critical issue.

Finally, lack of further study would be horrendous grist for the anti-vaccination bill. That's way we need to go on, and urgently I would add.

Dr. Brent:

Well, I have to preface my answer, which of course is yes. First of all, and I know others have said this, I have been very impressed with the presentations we had yesterday. It is not only the quality of the presentations, but I think the objectivity of the investigators. They really presented every aspect of the possibilities of it being a finding that is not causal versus one that was, and I think that is important.

As far as the answer to the question, I think it is not only one of further investigation, but what further investigation? With the birth defects, we have five areas that we look at when somebody alleges that something is possibly causal. One is what we discussed here today, Epidemiology: In our field it has to be consistency. In other words, we never depend on one epidemiological study because of what I mentioned yesterday, that if you look at enough T-tests, you are going to come up with a positive with relation to one birth defect. Therefore, you had better have that same birth defect come up in the next epidemiological study and the next one.

The second thing, the secular trend. I am impressed with the fact that some people here have information and believe that like the incidence of learning difficulties, behavior disorders and attention deficit is increasing in our population. I don't know whether it is or it isn't, but that kind of information you just can't throw around and say it's true or isn't true without data. And it is such an important area in our society. I mean it is the thing that makes a human being different from the other species, so it is such an important area of research.

The third area, one that we depend on a lot, is animal studies, and the fact is while you can't predict without knowing from a human study what you are interested in, the animal studies can be very helpful in looking at the mechanisms, thresholds and the incidence of the findings that you have in the human.

Then the fourth area is pharmacokinetics, which I think is crucial in this particular area, and the fifth is biological plausibility, and that's how we evaluate whether something causes birth defects or not.

So we are stuck with just Epidemiology here today and I think from the standpoint of further research, we need to find out whether these increasing dosages of methylmercury are really increasing dosages on the basis of pharmacokinetics or whether the fact is that each injection is a separate dose unrelated to the other one. I think that has to be done.

So what I wrote is that the results of the study that was presented reports a statistical association between vaccine exposure and certain neurological deficits. Two of the three explanations for the findings relate to patient selection problems and one explanation relates to exposure to the vaccine.

The incremental exposure to methylmercury. Statistical associations and causal connections are strengthened by obtaining the same results in a second or third epidemiological study, therefore, this should be pursued with appropriate populations.

Biological plausibility should be studied by performing pharmacokinetics in humans to determine the biological half life of ethylmercury in the blood levels of ethylmercury following administration. Appropriate

animals models utilize the ethylmercury in the threshold for neurobehavioral effects in each species should be initiated.

I think that is the basic data you would need to begin to get a handle on this problem.

Finally the implications are profound. I remember when I was an intern, I rotated to Boston and there was a woman there by the name of Pricilla White. Because I had been a researcher before I was an intern, she would come down and show me these placentas from mothers who were diabetic and because they were using DES, and she would say to me look at that placenta. Look how healthy it is from mothers who are on DES. Of course she was eventually crushed psychologically when they found out that it caused adenocarcinoma of the vagina. And the implications here are much vaguer. That was an epidemic which was horrendous. Causing learning disabilities and behavioral disorders. ADD is a tremendous problem in our society and I think it is one that we should be very concerned about.

Although my gut reaction, which is totally irrelevant, is that it probably is not causatic, the only way you can come to a conclusion is through the data, and that's the data I think we have got. Even if we put the vaccine in single vials and put no preservatives tomorrow, we still want the answer to this question. Because remember, epidemiological studies sometimes give us answers to problems that we didn't even know in the first place. Maybe from all this research we will come up with an answer for what causes learning disabilities, attention deficit disorders and other information. So I am very enthusiastic about pursuing the data and the research for solving this problem.

Finally, the thing that concerns me the most, those who know me, I have been a pin stick in the litigation community because of the nonsense of our litigious society. This will be a resource to our very busy plaintiff attorneys in this country when this information becomes available. They don't want valid data. At least that is my biased opinion. They want business and this could potentially be a lot of business.

Dr. Johnson: Thank you, Bob. I think you will agree that biologic research also needs confirmation, even when there is a hard Biochemical influence.

Dr. Brent: Absolutely.

Dr. Johnson: Okay, Loren?

Dr. Koller: In order to adequately answer question one, I took the prerogative to break it into two questions. The second one will answer your part of it.

Part one, is there a causal association between ethylmercury and neurological effects noted in the Vaccine Study Datalink project? The answer is no. Why not? From a toxicologists viewpoint there is no dose response relationship in some of the effects, particularly if you look at slide 18 where the cumulative mercury exposure, the rates for speech delay and ADD. At several doses those were lower or equivalent to the zero exposures for each one of those categories.

Another reason, in risk assessment the best human data is followed by the best animal data and it is used to determine no-L's and low-L's health affects.

Uncertainty increases in the direction from humans to animals, from high quality to low quality data or the lack

thereof. In my opinion the Seychelles study contains high quality human data, so that is the data that you use and it is supported even by the Faeroe studies and other studies in humans. The reason, there are 711 mother/infant child care, very few confounders. Children were exposed to high levels of mercury in utero, neonatal, during development of the nervous system, the most sensitive time. The children were vaccinated. There was continuous exposure throughout compared to single exposures in this situation. There were no adverse health effects in six neurobehavioral tests. As a matter of fact, in the higher group they scored higher on four of those six tests. Albeit, recognizing that there are other tests that may be more sensitive to detect neurological function. So therefore in my opinion there is no evidence that childhood vaccination would attain or exceed the Seychelles mean hair or blood mercury levels, let alone fourfold higher at the maximum range in that study.

So part two. Are the observations of sufficient concern to warrant further investigation? Answer is yes. Some of the neurological developmental disorders show a small, but suggestive increase in relative risk.

Dr. Johnson:

Loren, were you answering question two?

Dr. Koller:

No, I have question two as similar, but a little different.

Bob's first statement I think sort of laid it on the line. As you increase the vaccination, you increase effects, but you don't know. You have modified live viruses. You have different antigens. There is a lot of things in those vaccinations other than mercury, and we don't know whether this is a vaccination effect or a mercury effect. But I am almost sure it is not a mercury affect. Positive as a matter of fact, and there are several experts particularly

that have reviewed this, the methylmercury aspect who I think would agree with that due to dose response.

Dr. Johnson: Are you really comfortable with the way neurologic function was tested in the Seychelles?

Dr. Koller: I have to admit that there were many other tests that could have been conducted. In any of the mercury human exposures that have observed neurologic findings, most of them are negative clinically. We are talking about very subjective, very sensitive assays and yes, there could have been others done and there should be more done. That's part of number three. But we have to use the data that is available. If we went back to animal data, when you talk about suggestive and sensitive tests for neurological function in humans, it is much more difficult in rats and mice to detect those changes.

Dr. Johnson: Can't you put them out on those little floating pads and see if they swim and how fast they go through mazes?

Dr. Koller: You can, yes. In my opinion that is not quite as sensitive as it in humans.

Dr. Clarkson: Can I comment on that for a second? On the animal experiments. There is a lot of literature of animal data on methylmercury, and quite a lot on primates as well. The level of effect, the lowest effect level in those animals is about 100 times higher at least than what we are talking about in the Seychelles or the Faeroes or here.

Dr. Brent: But that is with methylmercury.

Dr. Clarkson: Yes, methylmercury.

Dr. Johnson: That's on a wave basis.

Dr. Clarkson: On brain levels, too. If you convert them to actual brain levels, you are talking about estimated brain levels of about 100 times higher. I agree that the animal data is useful in terms of mechanisms, in terms of what stage of gestation is important and so on, but I don't think that you are going to get human risk levels out of animal experiments. Because probably as you say the kind of tests you can do on an animal is not the same tests that you can do on a seven year old kid.

Dr. Johnson: Loren, if you are absolutely sure there is no causal relationship, why would you answer yes to question one?

Dr. Koller: Because I think there are other factors. There is many confounders that have not been evaluated. Biological and environmental. As a matter of fact, in question two one of my answers is there does appear, however, to be a weak positive association between increased numbers of vaccinations and some neurological endpoints. That is shown on slides 21, 23, 24 and 25. Because as you increase mercury, you increase vaccinations, so there could be several other factors in those vaccinations that are causing these effects.

There is also other types of vaccines that these children are exposed to. There might be a combination biological effect. It might be antigen effects. There is all kinds of possibilities here. Some of these are modified life viruses. I would assume they are modified life viruses. Something between the combinations or subsequent exposures in a sensitive population, or hypersensitive population may trigger some of these effects.

Dr. Clarkson: It will be interesting, Mr. Chairman, to know the conclusion of the aluminum meeting in Puerto Rico. What came out of that? Because we heard yesterday from the CIs that aluminum will correlate just as well as mercury

with these results. Is Dr. Myers here? What were the conclusions?

Dr. Myers: Well, first we didn't have this data to study. We didn't have available what we are discussing today. This study, so I am not sure.

Dr. Clarkson: What did they reveal about the all aluminum in terms of...

Dr. Myers: They thought there was an enormous margin of safety, that were well below concerns, but again they hadn't seen these associations. By summary we thought we were well below the mercury as well.

Dr. Stein: Well, of course I answered yes also, but first I want to say thank you to everyone for giving me a course in Epidemiology. I learned a lot. I want to also congratulate the group that did the study and the data analysis. It also gave me a great respect for the problems of evaluating vaccine safety beyond what I had ever known or expected before, and obviously I have been practicing pediatrics for a long time.

But I said yes because there are a lot of issues raised here. From my point of view as a clinician, it is not the subtle statistics that have been discussed and which are important, but really the endpoint. And that is the quality of these two diagnoses that have fallen out, attention deficit disorder and speech and language delay.

I recognize the limitations of a study like this, but I am going away uncertain whether these children, or most of these children, or a significant number of these children, really had ADHD or really have speech and language delay. I don't think the way the study was set up, even with the chart review, we really haven't been told about the quality of the diagnosis, the tests that were used to

substantiate these two diagnostic categories or the quality of the people doing the tests. In the area of neurobehavioral and neurodevelopmental problems, those factors are very important and it seems to me we are putting so much value on those outcomes without being able to substantiate. It is not like doing an SGOT where you can control for the quality pretty well. You can control for the quality of neurobehavioral and neurodevelopmental evaluations, but you have to have the knowledge to know how they were done, and we don't know that in this study.

Perhaps we could get better information by reviewing the charts in a different way, and for that reason I would vote yes. That we need to know about this, but I don't think you can make any conclusion that mercury is related to ADHD or speech and language problems in these children, given the lack of information about the quality of the diagnosis and that is your endpoint.

Dr. Brent: So what studies would you suggest?

Dr. Johnson: That's another whole thing. Dr. Rapin?

Dr. Rapin: I voted yes, but I had a question mark. I guess a yes question mark. I kept erasing and putting back in. I erased it finally this morning, but it was there. The question mark was because I was not at all convinced that the exposure level had reached a significant threshold effect after what I heard yesterday about mercury exposure.

Secondly, I was very impressed with the Faeroes and Seychelles, especially the Seychelles Island studies in which the children had much higher levels and no effect was detected.

I also felt that the study which we were provided on the 15 infants, five of whom were full term and 10 premature, was

strong because of the very smallness of the data set. So these are the reasons I have this question mark that kept wanting to come back.

In terms of why did I think we should pursue this. Well, as has been said by others, the first was the data that are there, they won't go away. They are going to be captured by the public and we had better make sure that (a) we counsel them very carefully and (b) that we pursue this because of the very important public health and public implications of the data.

I felt that the evidence, although statistically significant, the magnitude of the effect I thought was small and I was somewhat reassured by the chart review, and I really wanted to commend the reviewers because I have done a lot of chart reviews. It is a lot of data. But nonetheless, for reasons I will put in some of the later questions, I felt that the measures of attention deficit and language disorders and so on were weak. I have other criticisms that I put in the new methodology.

But again I want to thank you for this opportunity to review these data.

Dr. Johnson:

Thank you, Isabelle. I don't have anything substantive to add. I of course voted yes. There were two reasons. The stakes are very high and Bill Weil made this point. Any detrimental effect on infancy is serious enough to warrant as strong as possible efforts to define the relationship.

The second reason, as Bill also noted, despite of numerous efforts, and I agree with Bob Brent, I was impressed with the open mindedness and the concern in trying to ferret out what the relationship really was at all costs that was exhibited by particularly Tom, but also Phil Rhodes'

analysis, so in spite of that there is still some worries. So I put down for those two reasons a yes.

Any other comments on that question alone before we move on to question two?

Dr. Chen: Roger, do you want to read Alex Walker's?

Dr. Bernier: Alex Walker voted yes and he said if yes, why? You had a prior concern. You obtained mostly negative findings, but some positive results. If you do not treat this as a signal, other than much less responsible parties will do so, and follow up will be out of your control. Equally, the negative findings need to be pinned down and published. I think that is published. Need to be pinned. The negative findings need to be pinned down and published.

Dr. Johnson: That's very pragmatic.

Dr. Rapin: Can I make one comment about the business of the increasing prevalence of developmental disorders? I think that this parallels increasing education and sophistication of people who examine children. I can tell you from my own experience that 20 or 30 years ago I barely diagnosed autism unless it was so blatant that it stared me in the face, and now I see at least two new ones a week. And not so severe as the previous ones, so I think there is a tremendous change in the threshold of ascertainment. And yes, I have seen the California statistics which says it has increased 300 fold, but I would interested to know whether it has increased 300 fold in areas where there are physicians who have been trained in this recognition, as opposed to areas in which there are not.

Dr. Johnson: Thank you. We'll go to question two and go back in a reverse direction. The question is, if you think the observations on some specific neurologic developmental

disorders constitute a signal, how strong or weak do you consider the signal to be at this time, i.e. how much does the evidence support a causal relationship with Thimerosal containing vaccines?

I assumed this a number one. In my opinion the evidence today is insufficient to determine whether or not Thimerosal containing vaccines caused the neurological sequelae in question.

The diagnoses, even in the hands of experts, and the number of diagnoses are too easily influenced by variations in parental and physician sensitivity and concern, utilization of health care of similar merits.

The underlying biologic, toxicologic and pharmacologic data are too weak to offer guidance one way or the other. That is the biologic plausibility component of this, in my opinion, is too badly defined.

Now on the other hand, the data suggests that there is an association between mercury and the endpoints, ADHD, a well known disability, and speech delay as entered into the database.

Then here comes an opinion, well it is all is opinion, but it expresses a flavor, so I think it relates to what Dr. Bernier is trying to derive here. This association leads me to favor a recommendation that infants up to two years old not be immunized with Thimerosal containing vaccines if suitable alternative preparations are available. I do not believe the diagnoses justifies compensation in the Vaccine Compensation Program at this point.

I deal with causality, it seems pretty clear to me that the data are not sufficient one way or the other. My gut feeling? It worries me enough. Forgive this personal

comment, but I got called out at eight o'clock for an emergency call and my daughter-in-law delivered a son by C-section. Our first male in the line of the next generation, and I do not want that grandson to get a Thimerosal containing vaccine until we know better what is going on. It will probably take a long time. In the meantime, and I know there are probably implications for this internationally, but in the meanwhile I think I want that grandson to only be given Thimerosal-free vaccines.

Dr. Rapin:

I hesitated between a one and a two. I finally put in a two. My first statement was I thought there was an association, but it was not clearly causal. I felt that some of the things that made me feel this was weak was that children were counted as cases, irrespective of the age of diagnosis. As I said yesterday, many children who speak late turn out not to have language disorders, so there was no opportunity in the study for any change in diagnosis.

I felt the children were all studied below the age of six years and that attention deficit disorder, with or without hyperactivity, is an extremely weak diagnosis in pre-school children.

I felt that the diagnoses were made at different ages and the length of follow up varied, so that some children were only followed for a brief period of time. Those born in '96 and '97 were really seen for a very brief period of time.

I felt that even though some of the children were confirmed by referral to agencies for confirmation or for treatment by chart review, there was a lack of confirmatory tests.

I felt that the fact of parental worry on both detection and referral were important confounding variables.

I felt that the premature data which went in the opposite direction I found very troublesome, and finally the lack of family history data which would reflect on genetics, which I think are most important than environmental effects in all of the developmental disorders, was weak.

Dr. Johnson:

Thank you. Dr. Stein?

Dr. Stein:

Well, I also gave this a two to answer the question. My main reason was that the outcome measures of neurodevelopmental disorders do not provide enough specificity to make the diagnosis, as I said before. Again, we really don't know the quality of the diagnosis, and I will get into that in a moment.

Secondly, genetic influences were not considered. We need to know more about the family history, and when we get to the third question I will make a suggestion for that.

Three, there was a limitation. It occurred to me that the parents who take their kids for Hepatitis-B vaccine, especially in the early nineties when it was first recommended and at that time the conjugated HIB came out to give in early infancy. When was that? Ninety, so that's when infants received it for the first time because before that we were only giving it at 18 months, 24 months. These parents who knew more about the vaccine and might have accepted the vaccine may have been more sensitive parents and more sensitive to medical information in general, and more sensitive to developmental variations in their children. They may have raised more concerns during health supervision or well child visits, and requested evaluations for ADHD and speech delays.

Another aspect with regard to the introduction of Hepatitis-B is that when it was initially recommended by the bodies at the CDC and the American Academy of Pediatrics, many

pediatricians around the country were uncomfortable with that diagnosis because they had never seen a case of Hepatitis-B and wondered whether that was really an appropriate vaccine. And the question is were these pediatricians who gave the Hepatitis-B earlier more likely to be those who read more about it and also likely to be reading more about developmental delay and be more sensitive to that diagnosis. It is a hypothesis, but it certainly could affect the results.

Next, there is really no systematic review of the actual diagnosis of speech and language delays. I spoke a little to Tom at breakfast, and Isabelle has raised some of these questions about the maturational effects, particularly expressive language delay in boys at two to three years of age and how this can, in fact, be maturational in the majority of cases. Then when you evaluate them at four and five they don't have a speech defect. Eventually some may have learning disability pertaining to reading problems, but there is a lot of fluctuation to that diagnosis.

This is subjective. This is not the data that Tom pulled from the charts quantitatively, but many of the speech and language diagnoses were mild articulation defects, or articulation defects which are usually mild when they come from a general Pediatrician, and often reflect the greatest sensitivity of parents and concern and anxiety about parents, with what I would consider a developmental variation and not a true disorder. Whereas the speech pathologist may code it as an articulation deficit and give it a code, a diagnostic code. In fact, many of these are developmental variations. Just as the maturational expressive language delay in many cases at two and three years of age is a developmental variation. These aren't really disorders. And again, we are basing these results on these 1,533 children with speech and language disorders.

To answer your question which relates to this, Bob, about what tests would I use. There are certainly standardized tests to evaluate expressive and receptive language and articulation in early childhood, and certainly through the pre-school period. As well as for ADHD. There is standardized behavioral tests that can be used. With that in mind, it seems that at Group Health most of these children, if not all, were referred to a specialist. Or at least a Pediatrician who concentrated on working with children with ADHD specifically. Now we can assume, although we don't know, that person was really good at it and used standardized tests. On the other hand at Northern California Kaiser I am told, they don't have a specialized clinic for evaluation children for school under-achievement as a broad category and specifically for ADHD. So these children were probably diagnosed by Pediatricians, or perhaps in some cases a neurologist or a child psychiatrist. But again it is so heterogeneous we don't know the quality of that diagnosis as well.

Dr. Brent: Have you ever seen a child who has had that diagnosis who when you saw the child you refuted it or didn't support it?

Dr. Stein: Yes, many times.

Dr. Brent: My oldest son...

Dr. Stein: In fact there is some data on that.

Dr. Brent: My oldest son happens to be Chairman of Child Psychiatry at the University of Pittsburgh and he says about 25% of the children with that diagnosis do not have it, when they are fully evaluated.

Dr. Stein: Right, that would be what I was going to say, about one-third, and there are some data. I would agree with that, because ADHD is a diagnosis where the behavioral

symptoms overlap with a variety of other conditions, as well as with normal variation depending on the age, and that is another point.

The mean age of diagnosis of ADHD in this study was 49 months. Four years, one month of age. Well, ADHD is a very difficult diagnosis to make in the pre-school period. In fact, in our guidelines published by The Academy of Pediatrics, we limit the recommendation to the six to 12 year group because that is where most of the data is. There is very little firm data on the diagnosis of ADHD in the pre-school period. It certainly can be made, but in general it takes someone with lots of experience to do it, because so many of the behaviors of ADHD overlap with normal behaviors in this age group. Hyperactivity, impulsivity, inattentiveness. The developmental variation curve and the disorder curve really overlap tremendously and it takes a lot of experience to recognize it early on. Forty-nine weeks is very early.

Finally, and this is a question that was implied. Could the intake of fish by mothers who were breast feeding have influenced mercury levels in this study? We didn't look at the breast feeding issue. Now I assume from what you told us, it is known that mercury does excrete into the human breast milk. That is another very interesting factor I find that would need evaluation and further studies, but my main concern is again the endpoint. The quality of these diagnoses and all of our discussion is based on that. I think it kind of a weak foundation right now from what we know.

Dr. Johnson:

Dr. Koller?

Dr. Koller:

I gave it a one. First, as I indicated for question one, there does not appear to be a causal relationship between ethylmercury and neurological disorders.

Secondly, however, there is a weak positive association between increased numbers of vaccinations and some neurological endpoints.

Third, analysis of data has not included all confounders, including biological, environmental, as well as genetic differences.

And fourth, there is two to threefold differences in outcome repeated diagnosis between the two data sets, which is disturbing, the hyper diagnosis and interpretation of neurological disorders.

Dr. Johnson:

Dr. Brent?

Dr. Brent:

I personally want to congratulate Dr. Johnson on his grandson. I have a small series of 11 children, all who received the Thimerosal vaccine and they are all geniuses of course. But as Dr. Rapin points out, the genetics was probably most important.

Dr. Rapin:

My grandchildren are geniuses, too. I have two.

Dr. Brent:

Well, I circled one and I wrote the following. The epidemiological data is valid, as is the associations that were reported.

It is more difficult, if not impossible, to refute a causal association based on this study. Therefore, the question of causal association remains unanswered until we obtain the data that was suggested in the answer to the first question I wrote.

On the other hand, massive case control studies are sensitive, but frequently uncover non-causal associations, at least in our field. This would be much better if it were a 100,000 cohort study where you had controls and

exposures rather than a massive case control studies. You know it depends on what you pick as your controls, whether you end up with a positive association or not.

The most important information in the eyes of the epidemiologist is if the incremental exposure to the two categories of neurobehavioral effects that were likely to be effected, had increased relative risks. But when the pharmacokinetic data is evaluated, at least with regard to ethylmercury, the results may or may not support the incremental exposure.

Furthermore, the level of ethylmercury are in the range of mercury levels found in other populations as Dr. Koller referred to, where there are neurobehavioral findings and they didn't receive the vaccine.

Finally, the animal methylmercury data indicates a threshold for neurobehavioral events at a much higher level as mentioned before. This has to be determined for ethylmercury. So again, it is more data we need in other areas besides epidemiology.

By the way, I changed the question that I answered as has everybody else. The question I answered was, if you think the observations on some specific neurological developmental disorders to be valid, how strong or weak do you consider the data to be at this time? How much does the evidence support a causal relationship? I think that word "signal" bothered a lot of us because it gives you the feeling that you are talking about one piece of information and it was all the data that we looked out in those studies that we were evaluating.

Dr. Johnson: Nevertheless, in regard to causality you decided a one?

Dr. Brent: Yes.

Dr. Weil:

I put four and I did so for a number of reasons.

The number of dose related relationships are linear and statistically significant. You can play with this all you want. They are linear. They are statistically significant.

The positive relationships are those that one might expect from the Faeroe Islands studies. They are also related to those data we do have on experimental animal data and similar to the neurodevelopmental tox data on other substances, so that I think you can't accept that this is out of the ordinary. It isn't out of the ordinary.

The Seychelles Island studies, and somebody said the Faeroe Island studies both, were chronic exposures. We are not talking necessarily about chronic exposure. We are talking about a series of acute exposures and at one point in time that exposure is much greater on that one day than any of the Seychelles Islands.

The increased incidence of neurobehavioral problems in children in the past few decades is probably real. It may be a group of pediatricians, it may not be. I work in the school system where my effort is entirely in special education and I have to say that the number of kids getting help in special education is growing nationally and state by state at a rate we have not seen before. So there is some kind of an increase. We can argue about what it is due to.

Dr. Rapin:

Right.

Dr. Weil:

But there are certainly more kids with ADD and there are more kids with speech and language disorders than there have been in the past.

With regard to ADD I would only say that I don't think there is a diagnostic test. If you look DSM4, first of all

they don't even have criteria for kids under six. Second, they make a point that it is a label based on a constellation of findings and not a single test. The Conner's and all the other tests have been shown to have pretty different validity scores.

The symptoms, depending on whether you are a lumpner or a splitter. The splitters put ADD with every diagnosis where the symptoms occur. The lumpners say that if this kid has condition A and ADD, we will label it A. So there is a lot of variation among people who make this diagnosis, whether they are experts or not.

The rise in the frequency of neurobehavioral disorders, whether it is ascertainment or real, is not too bad. It is much too graphic. We don't see that kind of genetic change in 30 years.

There are also a number of toxic agents in the environment that could have done this. You see the evidence of Plopirophose as a neurodevelopmental toxic, and that has been widely used in the last 20 to 30 years as the most common household pesticide in the United States. I don't know how many hundreds of tons of this have been distributed over the rugs, carpets, dogs, cats and kids in our environment and it is finally being taken off the market as far as home use because it is a neurodevelopmental toxic. I think this relationship has to be investigated just as thoroughly as plopirophose was.

While the data are not sufficiently robust to accept a clear causal relationship, the difficulties in interpretation, the problems with alternative analyses and so on are not great enough to reject the possibility of a causal relationship. In other words I am saying it isn't there and I wouldn't give it a five or a six, but I don't think people would want to reject this and do so with the data at hand.

Dr. Johnson: You would neither accept nor reject, but you believe the data are not sufficient to accept or reject, but you would...

Dr. Weil: It is strong enough that I put a four.

Dr. Johnson: You assigned it a four.

Dr. Rodewald: What is the scale level?

Dr. Johnson: One to six. One is weak.

Dr. Johnson: Four is across the line. You are across the line toward the strong.

Dr. Caserta: Is the scale for how strong the signal is or how strong a causal association there is? That wasn't clear to me from the question.

Dr. Johnson: How strong the casual. This is causality. David Oakes.

Dr. Oakes: I want to pick up on something Dr. Brent said. I think this is a cohort study because you do have a defined population at the outset that you are following. There is a certain amount of fuzziness in the definition and incomplete follow up and obviously the differential ascertainment, but it is still basically a population that is defined at the outset that you are ascertaining outcomes. Imperfect, but you are ascertaining outcomes in a defined population.

Dr. Brent: I will let the epidemiologists answer y
you about that.

Dr. Oakes: So I think that should be put as a strength. We are kind of honing in on the weaknesses here, but that is a strength of it. That you do have the study in what is a pretty well defined population.

Dr. Snider: And retrospective cohorts.

Dr. Oakes: Retrospective cohorts, not prospective.

Dr. Rhodes: And that information is not figured out 10 years later. It happened in the past, but retrospective, perspective should refer to the information, not that happened 10 years ago or it is just happening today.

Dr. Oakes: With that preamble, I gave it a two. One of the strengths, doing the strengths first, is that it is in my view a cohort study. I find that somehow the first analysis is always in some ways a little bit more convincing than the reanalysis. Assuming the analyses were presented in the order in which they were done, the first analyses certainly showed some suggested trend tests, and I was very struck by the fact that you see a different pattern in relation to exposure from the neurologic diagnoses of interest than from the controlled diagnoses you chose, and I assume you didn't look at 27 others and reject those that didn't fit the hypothesis. I trust you did not do that.

So those are the strengths. The weaknesses. Clearly there is evidence of utilization bias and you presented a lot of evidence and a lot of discussion about that.

I did wonder why you didn't do more analyses which formerly included the potential confounding variables. You did have some visualization, number of shots or numbers of antigens or socioeconomic status. I don't remember seeing analyses where they were controlled for and you tried to look at the additional effect of the mercury. I think it is almost certain that you wouldn't see it, but I would like to see that analysis done. That may have caused a much more problematic issue in that there are going to be many other potential confounding factors that you do not have data on,

and you probably won't be able to get. At least certainly not on the entire cohort.

I don't think we have seen any evidence that the causal agent, if there is one, is Thimerosal and not some other constituent of the vaccine.

Dr. Brent: Could you say that again?

Dr. Oakes: We haven't seen any evidence that it is the mercury, if there is some damage being caused, that these associations are real, that it is an association with mercury. The question is what other things are in there that are also potential causal agents?

I am worried and I am not sure if it has been resolved or what the resolution is about the issue raised about the potentially unusual, possibly incorrect codings of some of the files and whether that really did have a very strong influence on the analysis. I'm not sure if that has been fully investigated or not.

Dr. Sinks: Could you...

Dr. Oakes: There was an issue that some of the codes may have looked very unusual for that time and may have been incorrect, and I am not sure whether the status of that is that they may have been incorrect or that they are known to be incorrect.

Dr. Rhodes: What we know about some of them, apparently it is information that was entered as it happened. It happened yesterday and the day before and it is being entered. There are certain quality checks on the data being entered, certain information, like what facility, what occurred, what manufacturer. There are vaccines that are entered a long time after because of various reasons and those do not always have those quality checks. Most of the ones that I

have seen as possibly being incorrect are these ones that apparently have been entered a long time after. Those are missing facilities, missing the manufacturer, so that leads me to believe that there is less quality control. I don't think there was ever a check in the program that said you can't possibly be getting a separate DTP because we don't do it. I don't think there were those kinds of quality checks.

Dr. Oakes: Are you in a position to say that some of the codings are definitely wrong?

Dr. Rhodes: We are in a position to say that some of them are very suspicious and need to be checked.

Dr. Oakes: But it is still a question mark at this point?

Dr. Rhodes: I can't say definitively.

Dr. Johnson: Thank you.

Dr. Clarkson: Well, Mr. Chairman, this is a historic moment. Two people from Rochester come up with the same number. I gave it two instead of one in a sense because I think speed delay is a plausible effect to a mercury compound in children or in infants. But I am very influenced by the pediatricians here who say for example speech delay is very poorly defined.

As far as the causality side itself is concerned, if you look at the mercury levels, those actually quoted in the reprints you have and those that we can calculate from what we know about the pharmacokinetics of methylmercury, these mercury levels, even given as a single shot, are still substantially lower than what you see in the Faeroes or the Seychelles, even though it is a single shot. I think this emphasized the need for this group to take a look at the pharmacokinetics in this study. I think it is something that can be done. You don't have all the body weights, but you

have the birth weight and you have the growth chart, so you can come up with reasonable numbers for body weights throughout the first six months. Then just take a look and see what these numbers are.

The ones reported in the literature are reasonable. Given the whole body weight involved here, given what we know about the pharmacokinetics of methylmercury, these numbers are reasonable. So I think it would be very helpful to come up with estimated blood levels here, to see how they relate both to the Faeroes and to the Seychelles. And I will reiterate that both the Faeroes and the Seychelles agree as far as postnatal exposure is concerned, there is no disagreement. Both studies have not been able to find anything connected with postnatal exposure in infants. So I disagree a little with my colleague down here because he mentioned prenatal data, but the postnatal data, which you were concerned here with postnatal exposure, is consistent in this respect.

Finally, I think there is some evidence that there is a confounder here. If you look at the correlations for cumulative exposure at one month, if I read this correctly, Tom, you were finding correlations with language and speech delay at the one month. To me the increasing mercury levels in your population at one month due to 12.5 micrograms, is so small that it would suggest to me that you have a confounder here. That this is not due to mercury. The increase in a kid of 3.3 kilograms with 12.5 is within the normal range. It is hardly detectable. So this suggests to me that if you do get a correlation here, it is probably due to other confounders or other causes here. There may be a mercury effect, but it suggests here that there are other effects that would explain it. As you yourself mentioned, that the first cause is the parents attitude. I agree with that.

Dr. Brent: Dr. Clarkson, could I ask you to elaborate this point you made about postnatal exposure in the Faeroes and the Seychelles yield the same result. I would like to understand that.

Dr. Clarkson: The same result is that we didn't find anything.

Dr. Brent: What postnatal exposures?

Dr. Clarkson: The postnatal exposures in the Faeroes were levels in the children at 12 months of age were correlated with neurodevelopmental outcomes. Actually in the Faeroes the paper is about the 1996 paper. It is the same cohort. This is the cohort where they found prenatal effects at seven years of age. Now in that same cohort at 12 months of age, a comparison was made with levels in those kids at 12 months of age. Not in the mother. Not the prenatal levels, the postnatal levels at 12 months of age. In that report, no adverse effects correlated with these postnatal exposures. Are you with me?

Dr. Brent: And that is controlled in some way for the prenatal exposures?

Dr. Clarkson: No.

Dr. Brent: That is what I am confused on. You have got two different...

Dr. Clarkson: The difference are no correlation with the postnatal exposure at 12 months of age. Now the prenatal, there was an effect of prenatal exposure, but that effect was picked up at seven years of age. So in the Faeroes study which is the only one that found a prenatal exposure, they could not find any postnatal correlation, nor could we in the Seychelles. We looked at kids at six months of age and 19 months. We couldn't find any correlation with postnatal.

Dr. Brent: Because that is a crucial point. Everyone keeps talking about the difference between the study that found positive results and the study that found negative results. You are saying in fact the studies have both found a negative result for postnatal exposure. That is crucial.

Dr. Clarkson: They did not find anything. If you find something, perhaps sometimes people say that's a positive result. So we have to be clear about this. Both in the Faeroes study and the Seychelles, they were not able to find any correlations between measured postnatal exposure and the outcomes.

Dr. Brent: What age again?

Dr. Clarkson: Twelve months, and the outcomes in the Faeroes was attainment of the classic developmental milestone.

Dr. Brent: But at age seven?

Dr. Clarkson: At seven years of age there was a correlation between neurobehavioral effects and prenatal exposure, and there was no correlation at seven years with postnatal exposure.

Dr. Brent: Postnatal, and the neurologic exams in the Seychelles were done?

Dr. Clarkson: At six months, 19 months, 29 months and 6 years.

Dr. Weil: There is also a 96 month.

Dr. Clarkson: We haven't published that yet.

Dr. Weil: But there is one?

Dr. Clarkson: There is one, yes. They are in the hands of the statisticians. They are physically doing it.

Dr. Weil: But it is also very difficult to determine the postnatal exposure levels, because nobody measured how much mercury they were taking in every day for seven years.

Dr. Clarkson: That is correct.

Dr. Weil: So the postnatal data is very worrisome in terms of what the actual exposure was. In addition, the sensitivity of the evaluation is not what we would have hoped in terms of if we do these kind of data before that had happened, we might of looked at somewhat different.

Dr. Clarkson: The problem with the Faeroes for instance is that they were getting actually beneficial effects. So that in terms of the attainment of the classic milestones of development, these were attained more rapidly the higher the mercury level at 12 months of age. I think they gave a very plausible explanation for this. That there was a confounder and that confounder was breast feeding. They showed the longer the breast feeding period is, the higher the mercury levels, and the well know literature that breast feeding is good for you. So this seemed to be a very reasonable confounder.

Now I don't see you are going to change that picture by any other kind of outcome. These kids were doing better.

Dr. Weil: Well, they were doing better in terms of development milestones.

Dr. Clarkson: Right, do you think they would do worse neurophysiologically?

Dr. Weil: I don't know. I have taken a lot of histories of kids who are in trouble at school. The history is that developmental milestones were normal or advanced and they can't read at second grade, they can't write at third grade, they can't do math in the fourth grade and it has no relationship as far as

I can tell to the history we get of the developmental milestones. So I think this is a very crude thing as a measure of neurodevelopment. Hopefully we will be looking at much more sophisticated measures of neurodevelopment the next time we get into this kind of situation, but I think those of us who work with kids with neurodevelopmental problems at school age would say that there appears to be very little relationship, except the severely mentally retarded and so on, between those kinds of things we are concerned about. Learning disabilities, reading disabilities, visual perceptual disabilities and developmental milestones.

Dr. Rapin: Most developmental milestones. Most developmental milestones, but not language.

Dr. Weil: But most of the measurements that pediatricians make for developmental milestones are motive.

Dr. Stein: But those are historical milestones you are getting from parents of children who are school age, so you are dealing with memory at that time. That's the problem.

Dr. Clarkson: This by the way was the Faeroes. In the Seychelles we didn't do that. Well, we did milestones. What we did is Fagen's test and we did the Bailey's, so the outcome measured in the Seychelles were different.

Dr. Weil: And again, we could argue for hours about that, but I won't do that.

Dr. Johnson: Kevin?

Dr. Sullivan: I gave the value as one. I think the strength of the associations are mostly weak and the weaker the associations, the more likely bias might explain some of this. The issues on biologic plausibility, it seems about a

maybe. The dose seems to be small. There seems to be some issues of whether these small doses could cause these effects.

Dose response. There does seem to be there may be somewhat of a dose response with some of the outcomes.

One issue would be the quality of the data. Using observational data, computerized data sets. This is not designed as a study to look at the effects of these vaccines on these different outcomes, but it is using data collected for other reasons, so it is not a carefully controlled prospective cohort study to study. We are using data that is really collected for other purposes. That is not to say that the VSD, I think it has been extremely useful. You could probably look at some of these associations with a large sample size. I think it has been very useful for that. I think always in the back of our minds we have to remember that anything you can find in this has to be interpreted very cautiously because of the way the data are collected.

One issue is the outcome. We have a lot of experts here in the area. That they are poorly defined. No consistent diagnostic criteria applied, and with probably a lot of misclassifications. Some who are called as having this diagnosis may not have had it. There were a lot of children who were not given this diagnosis, and maybe they did. I am not sure which way that misclassification works. Differential or non-differential according to the vaccines. I don't know, but we know there is a lot of misclassification probably in the outcomes.

Exposure to the vaccine. We really haven't talked about that too much, although some information was given that there is a misclassification on vaccines. That some children whose record may say they have been vaccinated when in fact they have not received that vaccine, and some

of the no vaccine individuals may actually have received a vaccine. So we have a misclassification of exposure.

Another issue is, is that a differential misclassification or a non-differential. It may be if the parents were getting the children immunized early are being more observant of the child's development and growth, so that made me think that there may be a differential misclassification of the outcomes.

Talking about some of the analyses. Well, there were a lot of statistical tests. I think we have to be concerned that by chance some of those might be due to chance alone. So we cannot always look at the P-value and say every one of those is true. If it is not statistically significant, it's not true. I think there has to be a lot of caution in there.

One thing that was not brought up was the assessment of effect modification. I always feel that if you are going to control for something, that you really should look to see whether there is an effect modifier of the relationship first, because you don't want to control for something that modifies the vaccine or mercury levels in the outcomes. I never saw any information. Looking for effect modification might be interesting. There might be subgroups of individuals where maybe there might be some stronger association and no association in other subgroups.

Again, part of the VSD, there is lack of some of the variables that might be useful for assessing or that might modify this relationship or confound it. There was information given on birth weight, a very small sub-sample. FCS is not known very well. Ethnicity, breast feeding, so there is a lot of things that may be somehow involved in this that we really don't have good, solid information on that.

As far as the mercury levels, again I think it has been brought up that we are talking about kids who are getting challenges with lots of different antigens, the more the mercury exposure is going hand in hand with the number of injections and other exposures, so in general I think it is a weak association from the evidence we have seen here. There are lots of problems here, but I feel we should probably go on and look at this a little more carefully.

Dr. Johnson:

Thank you. Paul.

Dr. Stehr-Green:

I also gave it a two. The evidence for causality is sparse because the determination of causality is based on many factors, not just statistical association and how strong that association might be, and many of my reasons have been stated already.

I sort of went through and weighted for and against. Temporal association. I think there is evidence of temporal association in only the barest sense that I think occurred before diagnosis. However, there was nothing to show that the distribution of those outcomes, indeed they are real because I have a lot of questions about the consistency and voracity of those diagnoses. There was no analysis to show that the distribution of those over time is nothing different from the normal background breaks of occurrence.

In terms of strength of association, even though I think there was evidence to form an association, I think at best they demonstrate a weak elevated risks for some of these outcomes.

Consistency with other findings. There really are no other findings of similarly designed or similarly focused studies, at least of which I am aware of or at least that was presented, so we can't really say that this is consistent with other findings.

Biological plausibility for the reasons that Dr. Koller stated, the levels of exposure in this study were likely lower than exposure levels seen in other studies where no effect was observed, so that kind of mitigates against biological plausibility.

Although in effect it was an extrapolation from methylmercury to ethylmercury raised uncertainties. Even so, in the balance there is not tremendous evidence about plausibility.

Dose response. That was really the hardest statistical analyses that was presented, and I think some of the analyses demonstrated dose response curves at some age levels. But again, the inclusion of these supposedly low exposure groups, the whole question of plausibility of ascertainment I think has to be weighed when considered against the relatively small significant dose response curves calls those into question.

Finally, the issue of reproducibility, which is related to the issue of consistency with other findings. We never will be able to do human experiments per say, but there may be opportunities to do other types of studies as a dimensional rating that we will get into in the third question. We may be able to look at this more carefully to see if we can reproduce these, using operational data of course, to reproduce these effects.

As an editorial note, I think asking us to assess causality was kind of a foregone conclusion. There is no way we are going to find that this was a causal relationship, based on the data and evidence presented. So I am not sure in that respect the results will be useful because I am not sure there was ever any possibility that we are going to find other ones.

Dr. Johnson:

This is Dr. Walker's comments. He gave it a two in favor of causality. Stronger results which validated data. As the data were validated, the results got stronger, at least in some cases. Relations do not depend on the extreme values of vaccination status.

Against, uncertainty about the clinic. Confounding.

Second, plausibility of medical, social artifacts and alternative explanations.

Third, lack of supportive or event related toxicology, pharmacology.

So he leans for a 1.8.

Okay, we are a little bit behind on the previous schedule and we have tightened the schedule up by 30 minutes, so what Dr. Bernier has asked is that we take a shorter break than allocated. I think we trying to end at noon, is that right? I think we will try to leave at noon. My feeling is that the research can be shorter than this last round, is that the feeling? A lot has already been covered, plus it is written down.

BREAK

Dr. Chen:

We felt that it was important to bring this data to wider scrutiny despite it being only phase I and despite as someone argued, that the data has shown very low relative risk.

The main reason for that, I think we felt that unlike most other vaccine safety signals in the past which have come from VAERS and despite the problems of the events about the VSD, that in general the database was designed to look at safety issues and give them the precision on the exposure

side. We felt that it was really a hard quality of initial data source. That the dose response was probably some, but not all, selected biological plausible outcomes that may be associated with mercury, and that while we were concerned a bit about the multiple comparison issue, it is hard to explain away a dose response curve based on those multiple comparison arguments. And that whenever we tried to tier the data in terms of increasing the specificity of the diagnosis, in general we found either consistent or a higher relative risk. This was even when we tried to restrict it to more than one visit and when we did chart reviews and in general, in epidemiology that suggests that it is not a finding.

We were very much considered about the utilization bias, as well as the lower level exposure groups, but that when we picked this non-biologically plausible outcomes, in general they came up with different curves. So that led us to kind of think those other biases should be consistent throughout, and we definitely felt that more definitive studies were needed with systematic review and Frank and Bob will present that. But that over 10 years of working with this database with probably over 25 studies over time, these very experienced PIs were worried that this information, given the current climate, do warrant a greater scrutiny other than us just plodding along, finishing our cases, et cetera.

Again, I wanted to thank you all and give you all the basis.

Dr. Stein: May I ask you, 25 other studies came from this database?

Dr. Chen: Yes, and we have some review papers that we will supply you.

Dr. DeStefano: I am going to start out with what is supposed to be the next step. What I am going to do is basically try to summarize

and give you a little bit about what we have going on in the next step. Summarize what has been suggested by many of you as the next step, and then turn it over to Bob Davis, who as you recall when I talked yesterday, he said this was like a protocol. This was going to be at least a two phase study. The first phase was a screening and that is what we have been discussing, but the second phase was going to be the more definitive study. So I will turn it over to Bob and see what he proposed for a phase II study should be, now after the things we have discussed.

First of all, the next step to a possible association. I think we mentioned this yesterday in terms of the consistency of findings via the replicator. We are trying to replicate these with data for another HMO. We have been in contact with Harvard Pilgrim Health Plan in Boston and they are trying to put together a data set similar to what we had in the VSD so we can try to replicate these analyses in other populations. They expect to have about 20,000 to 30,000 children. This is on the order of the Group Health size cohort. They will try to use the same methods as VSD, although here we will have more a restricted A priority hypothesis if you will. Our intent is primarily to look at the speech billings and attention deficit problems. And we both have put it on here results by 21.

Dr. Chen:

Is it possible for CPP?

Dr. DeStefano:

Yeah, a suggestion was made about CPP. I am not familiar with the data set. I have some questions about that, if you can fill us in. What ages were these children followed? Would they have been seen for these kinds of problems and the vaccination...

Dr. Rapin:

Yes, through age seven. Talk to Karen Nelson and she will tell you all about it.

Dr. DeStefano: Vaccination data?

Dr. Rapin: But you have to talk with Karen Nelson because she has been minding this database for developmental problems for years.

Dr. DeStefano: Vaccination data.

Dr. Brent: The original collaborative perinatal project was to look at the cause for cerebral palsy. That is what it was originally designed for, and mental retardation. They accumulated everything. They registered the people at the time they became pregnant, so they had all the information collected on their prenatal care and they had many visits before their babies were even born.

Dr. Rapin: The last visit was at age seven.

Dr. Sinks: Just to follow up on that, the National Institute for Child Health and Development, I don't know if they have vaccine information.

Dr. Orenstein: They do, they have published data on seizures. I know on whole cell DTP out of that database.

Dr. DeStefano: There has been much discussion about a study of exposure. If you talk to Michael Gerber, you can see that there is one study in progress that NIH has been doing at the University of Rochester. There is no data available.

Dr. Clarkson: The data samples of urine and blood from the infant and the hair samples from the mother...

Dr. Gerber: Do you want to just describe the study? What it is, Tom?

Dr. Clarkson: Would you like to?

Dr. Gerber:

Well, this is the attempt to look at the pharmacokinetics of ethylmercury in 40 infant/mother pairs. What we are attempting to do is get one group of infants who were not exposed to Thimerosal containing vaccines. It turns out fortuitously that the Bethesda Naval Hospital has not been using Thimerosal containing vaccines for the past two years. So we are going to use those infants, and a group of infants from Rochester, some of whom who were exposed to large amounts of Thimerosal and others who were exposed to a moderate amount. The idea is to look at these infants' blood levels, urine and stool within one month of having received vaccination. Then at the same time look at maternal hair samples, as well as dietary histories from the mother to get some idea of potential baseline exposure in utero. Then get some sense of the pharmacokinetics of ethylmercury in these patients.

Dr. DeStefano:

There was a suggestion made earlier it is important in these pharmacokinetic studies that humans, if they would just adequately address this concern. And it was also suggested that we do more animal studies. One or more studies in animals.

Dr. Weil:

Just let me paraphrase that. That has to be neurodevelopmental toxicity studies. When you talk about animal studies, there are millions of kinds of animal studies, but there are now specific guidelines for neurodevelopmental toxicity and that is what you need to be looked at in this particular situation.

Dr. Myers:

And I would suggest that they ought to be ethyl versus methyl as well, to distinguish the relevant contribution. Another study which I think Dr. Clarkson is doing is looking at the contribution of ethylmercury and the types of vaccines that were given to children in the Seychelles. To look and see if we can

determine what the contribution of ethylmercury to their exposure was.

Dr. Clarkson: I knew you would say that to get it in the record, but we are doing our best to find out about that. About the toxin exposure.

Dr. Braun: Just to mention, and I think you may have seen the protocols, Frank, there are some collaborative studies planned between the Center for Biologics at FDA and NIEHS, looking at animals and the pharmacokinetics and also, if I remember correctly, histopathology in experimental animals dosed at various ranges of doses of ethyl and methylmercury.

Dr. Stehr-Green: Frank, when you were describing the study of the NIH Bethesda study, University of Rochester, it seemed to be a very valuable, natural experiment source. If Bethesda has been giving vaccines without Thimerosal, is it possible to look at some of these same health outcomes? Do chart reviews? Or does the data exist in some way? That way you could separate out the other vaccine component effects from the Thimerosal effect.

Dr. Brent: But they are only two years into the project. They wouldn't have children old enough.

Dr. Gerber: Yes, according to last year's. I don't know how big that cohort is. It is the Bethesda Naval Hospital's Pediatric Ambulatory Clinic.

Dr. Stehr-Green: It may only be a future potential at best.

Dr. Gerber: It could, yes.

Dr. Rodewald: I was wondering, when you are talking about the research agenda if it would be helpful to pose it in questions rather

than in types of studies and things like that. What is the hypothesis that you want to test. We have done it on a couple of those, but when it just says further animal studies, that is rather vague.

Dr. DeStefano: These are just notes I have taken on the discussion.

Dr. Rodewald: But I think that may be the helpful discussion and say what questions? Part of what sounds like it was discussed is the impact. We really tried to address causality directly and I wonder if that is something that is going to come up on a future slide in here. Because I am not sure how well you are going to be able to hit at some of the causality questions in here. I think to gradually try to hone in on that would help.

Dr. Brent: With regard to sort of the administrative problems here, I can understand that with regard to the epidemiological studies, your group would be involved in orchestrating in a positive sense. Orchestrating the epidemiological data that is available in the United States. But with regard to the animal studies, who would be responsible? Would it be the FDA, because they have a wonderful facility in Arkansas with hundreds of thousands of animals and they could put together a valid project. Maybe you would want input from the group here to tell them exactly what you would like.

And the pharmacokinetics, who would do that? Who would have that responsibility, because that is a small study to look at the mercury pharmacokinetics in a small population to get an idea of how long it lasts and what would happen after five doses? Would you have any different blood levels? In other words, we need some kind of administrative input in order to have all these things going on at the same time.

And I say that because I wrote a last paragraph. It is sort of frightening to me, but I will read it. By the way, I have been involved in three lawsuits for the vaccine group, and they happen to be people who were given vaccines who were pregnant and the allegation was that the vaccine caused the birth defects. Let me tell you, if you want to see junk science, look at those cases. It is amazing who you can find to come and testify that such and such is due to a measles vaccine. They are horrendous. But the fact is those scientists are out here in the United States. So let me read what I said.

The medical/legal findings in this study, causal or not, are horrendous and therefore it is important that the suggested epidemiological, pharmacokinetic and animal studies be performed. If an allegation was made that a child's neurobehavioral findings were caused by Thimerosal containing vaccines, you could readily find a junk scientist who would support the claim with "a reasonable degree of certainty". But you will not find a scientist with any integrity who would say the refuse with the data that is available. And that is true. So we are in a bad position from the standpoint of defending any lawsuits if they were initiated and I am concerned.

Dr. Weil:

So it may not be the government doing some of these studies. If you could use any of the precedent from other drugs and other chemicals is smaller than the fact of dumping this back on the industry that uses the vaccines and ask the company to produce these studies. That has certainly been a pattern for an awful lot of things.

Dr. Johnson:

Bill, when you say fund the studies, is that what you meant?

Dr. Weil:

Well, some of the companies will do them in-house because they have the expertise. Others may fund

somebody else to do them, depending on the amount of expertise. But the government has had a tendency, and I don't know if they will in this case, but to rely on the industry to deal with the basics and then neurodevelopmental studies. With a little pressure, they may change their minds, but I don't know that.

Dr. Sinks: Just to perhaps answer Dr. Brent's question about the part of government that may be responsible, the National Toxicology Program at NTP is an inter-agency, collective if you will that is basically housed at NIEHS, and I believe Miles Braun probably was referring to a collaboration within NTP, which has FDA as part of the executive board, CDC, ASTDR, NCI, and I am pretty certain that George Lucere, who is about to retire at the end of the month, but was involved and I think that they are doing some initial bio assays, either in Arkansas now to look at ethylmercury. I think that is an appropriate route to be talking about.

Dr. Brent: I agree, but what bio assays?

Dr. Sinks: I am not exactly certain what they decided to do. I think Miles probably described it.

Dr. Braun: I looked at the protocols and I can't really quote them to you. I think it is important to underline that these are planned studies and they depend on an argument, at this point as I understand it, between different agencies. NIH and FDA, but I think coming out of this meeting, if it is felt that is an important project to carry out, then that certainly could help it actually coming into place. Don't get me wrong, these are not underway. They are planned. They have protocols. There is a lot of thought that has gone into these, but that is about as far, as I understand, where they are.

Dr. Brent: I think it is very important if that group, which is an excellent group, is planning to study, if they have some consultative help from some of the people who are here because we now have heard all the information here and have a wealth of information. They could provide the animal experimental people valuable information while planning the project. I would hate for them to go through with a \$50,000, \$100,000 or a \$200,000 project and not have had information from this group which would help them design a better study.

Dr. Braun: Well, anybody who would like to contact the people who are investigating, I will be glad to pass this on. If you want to give me a card or something, I will be glad to pass this on to those people who are planning on carrying this out. I would think they would want to get the kind of consultation you are talking about.

Dr. Chen: Maybe arrange for some senior body with the protocols to be sent to the consultants for review. Be it us or be it Tom's group or whatever.

Dr. Myers: I think in answer to you question, although the meeting has been convened and is being led by CDC, if you look around the room, we are all here from each of the different agencies, and the reason for that is we are looking for the input for cross agencies, not just for CDC.

Can I go back to the core issue about the research? My own concern, and a couple of you said it, there is an association between vaccines and outcome that worries both parents and pediatricians. We don't really know what that outcome is, but it is one that worries us and there is an association with vaccines. We keep jumping back to Thimerosal, but a number of us are concerned that Thimerosal may be less likely than some of the other potential associations that have been made.

Some of the other potential associations are number of injections, number of antigens, other additives. We mentioned aluminum and I mentioned yesterday aluminum and mercury. Antipyretics and analgesics are better utilized when vaccines are given. And then everybody has mentioned all of the ones we can't think about in this quick time period that are a part of this association, and yet all the questions I hear we are asking have to do with Thimerosal. My concern is we need to ask the questions about the other potential associations, because we are going to the Thimerosal-free vaccine. If many of us don't think that is a plausible association because of the levels and so on, then we are missing looking for the association that may be the important one.

I thought I would put that out. That we shouldn't just think in terms of mercury.

- Dr. Gerber: Just to follow up on Marty's comment, it seems to me that during the time that this study was done, 1992 to 1997, at least at Northern California Kaiser, there was a substantial number of children involved in vaccine trials. The vaccines that those children would have received would not have shown up in the CPT coding. When you go back and reanalyze the data, I wonder if there would be some way you could determine what other vaccines these children may have received as part of the clinical trials?
- Dr. DeStefano: We know if they received any experimental drugs.
- Dr. Gerber: You would know?
- Dr. DeStefano: Yes, we have the data.
- Dr. Gerber: Okay, but you did not include that in the analysis?
- Dr. DeStefano: Tom, did you look at those?

Dr. Sinks: I haven't separated them up.

Dr. Rhodes: As far as we knew none contained Thimerosal, so they were not included.

Dr. Gerber: Right, but it may not be Thimerosal.

Dr. Chen: To address Marty, I think that is quite reasonable, although we have a limited amount of manpower because of what we just studied. At the moment, I would think most people around the room would argue these are biologically plausible outcomes potentially related to mercury, and then we will keep the other ones in mind. But hopefully we could do some of these studies to kind of rule it out, and yet if the association still stands, then we can look at some of these other hypotheses. That is the first step. Given the amount of time today, maybe just focus on mercury.

Dr. Myers: I agree with you, Bob, but the think the conclusion is there is an association between vaccines and the outcomes that we cannot reject and of which one compliment of the vaccines that is associated is Thimerosal, but it is only one of the associations. I don't think it is any more plausible than some of the others. And I think I heard several of the consultants say the same thing.

Dr. Johnson: That is an important prospectus, but our charge today is to focus and pick out obviously the mercury and focus in on that. That is a pretty tall order.

Dr. White: I thought we were looking at future studies and how to delineate what is causing this. If they gave it a one and a two, they thought it was a causality in this and there is aluminum. You could run these tests in another arm, in an animal study, a lot cheaper than restarting it up again. I think it is a good suggestion and the industry representatives that provide bulk for these vaccines. I'm

glad you invited us here because I think they would be willing to work and provide that. It would be cheaper to add that arm.

Dr. Caserta:

One of the things I learned at the Aluminum Conference in Puerto Rico that was tied into the metal lines in biology and medicine that I never really understood before, is the interactive effect of different ions and different metals when they are together in the same organism. It is not the same as when they are alone, and I think it would be foolish for us not to include aluminum as part of our thinking with this.

Dr. Stehr-Green:

I think generically, you know there are books on mercury and Thimerosal. Because of these other concerns, I think it will be important when we design all of these studies to think about ways of excluding other possible genealogic agents, either in the design or in some way so they can do the analysis that way.

Dr. Brent:

The advantage of the perinatal project is some of the vaccines that would be included today were not available then. The only thing we had as far as I recall is the Diphtheria, Tetanus and Whooping Cough. You didn't have the Hepatitis. You didn't have some of the other vaccines, so that is a unique group of people that could sort of sort out some of the other issues that have been raised.

Dr. Orenstein:

How about smallpox?

Dr. Brent:

They had smallpox.

Dr. Orenstein:

You have to add smallpox and IPV. In fact, one of the studies from the perinatal project suggested an increased risk of tumors in the off spring of parents who received three CBL. Heard of these associations.

- Dr. Sullivan: Are there any clinical trials begun in the last 12 years where it will be enough variation, like in HIB trials for example, where Thimerosal containing vaccine was given to some and not to others in a population that might hopefully define information about developmental disorders later on? For example, the HIB study that was done at Kaiser. I don't know how varied the Thimerosal exposure would be in those kinds of studies.
- Dr. DeStefano: In Northern California we tried.
- Dr. White: There was a huge study that was done for pneumococcal conjugates and as a control they used the meningococcal disease and I don't know, it's either neither contained Thimerosal. Well, there you go.
- Dr. Myers: That's right, and also we would have to wait some time before, but that original HIB efficacy trial, California used the single dose vaccine that did not contain Thimerosal.
- Dr. Verstraeten: It is still interesting because it contained some of the other.
- Dr. Rhodes: One thing that hasn't really come up is there are plenty of other kids, even just at NCK and Group Health, who haven't taken part in the current analyses. In other words, if you look at current eight, nine and ten years old, and if you had some information about what their Thimerosal containing vaccines might have been when they were infants, and if there would be enough variation in those kids and it's a controlled setting, you are looking at outcomes maybe you would feel more secure about than the seven, eight, nine, ten. There are kids at those ages now, but the question would be how good would the vaccination information be on Thimerosal going back to that same time. If you had that information you wouldn't have to wait three of four years.

Dr. Weil: Most of those kids are in schools that require a vaccination record, which includes not only the date and the vaccine, but the lot number, so if you look at eight and nine year olds now, you will find probably out of most school systems, some pretty good immunization records. My guess is it would not be hard to find a sample.

Dr. Orenstein: Not only that, but the validity of some of those school records has been problematic in terms of people getting extra vaccinations because they had vaccinations that did not get reported into the school records. Minnesota turned out to be very accurate. Dallas County turned out to have a substantial inaccuracy of data.

Dr. Rhodes: I'm talking about kids who are in the HMO that have reflected of this other data.

Dr. Rapin: How about the Mayo data?

Dr. Chen: Why don't we let Bob and Frank present exactly the sense this cohort has the best information and exposure going back to about 1990, and so be able to kind of quickly finish.

Dr. DeStefano: I think one last thing was going to go like, what would you do in that kind of follow up study? I think the same issues would come up during Bob's presentation. He is going to present us a more specific proposal rather than general issues.

Dr. Davis: Probably not as specific as you had hoped.

Dr. DeStefano: Not as general as I have.

Dr. Davis: As we have all talked about, current studies lack a lot of data, including mercury intake of the mothers during pregnancy. I am talking about the current studies that we

are looking at today and yesterday. And also a lack of information on breast feeding as

Dr. Clarkson pointed out, which is breast feeding is a mercury exposure vehicle and also a way to improve neurologic functioning.

It is not obvious how this might affect the current study. It simply is not obvious to me, could it be related to both the outcome which is very plausible, but could it also be related to Thimerosal exposure at one to three months? That is tenuous, but I am still not convinced. I think Phil made the strongest argument that there might be some confounding that has actually entered into our data. We thought this was actually a wonderful, natural experiment when we started out. Phil pointed out the fact that it is a natural experiment, however, it may not be wonderful.

Next slide. Just to point out very quickly that these current studies also lack the usual suspects, which are alcohol, smoking, nutrition prenatally, lead exposure and nutrition postnatally, demographics including race and ethnicity and socioeconomic status.

Again, while it is clear that these are related outcomes, neurodevelopmental and neuropsychiatric developments at five, six and seven years of age, it is not clear how these are related to Thimerosal exposure going to three months of age.

Next slide. This is the thing we are all worried about. Due to time, I am not going to go into it again. It's a signal that has remained after taking birth weight into account, although crudely, after we have limited it to kids who have had at least two visits for the outcomes of interest. When we limited it to second diagnoses, and then when we excluded children with competing cognizant diseases of interest, this signal has remained.

Next slide. I am going to be a little controversial here, in that I think there is a possibility to conclude the analysis without going a lot further. Phil pointed out something that he went through very quickly and then I spent a lot of time thinking about that last night, which is that in our zero exposure group we have a lot of kids that were just about to be vaccinated, so we may have been too conservative in how we considered our zero vaccination groups. So I think we should play around a little bit with widening our vaccine exposure window. I am not talking about the dosage, I am talking about our one month time period. I think a small group of us should sit down and think about perhaps at one month and play around with the definition there. That is worth revisiting because I was worried, as Phil pointed out, that our one month window excluded kids who were literally one or two days away from being vaccinated. There was other data that I won't get into now that actually suggested that in fact may have played an important role.

Also this business of our stratification by time. I think we have beaten this one into the ground. I think we may have dropped a lot of risk sets if we stratified by time by one month. I think we should go back and reconsider using two, three or four month time windows. I am not a big fan of secular trends occurring that fast within this time window. I think we should look at that.

I think we could perhaps include some previously excluded children, but this is something that Miles Braun, I, and some other people were talking about, which was using another controlled outcome. Not gastroenteritis, not conjunctivitis. I think we should another control outcome group, chronic abdominal pain, which I think all the pediatricians and parents in the group would realize that parents who are likely to bring their children in early for vaccination would also be more likely to bring this in for

medical attention, and then see if the same signal persists with recurrent abdominal pain.

Not so much for this one because the signal disappears with these three reanalyses. I actually think we should stop. I'm not on your committees. I'll say that being very controversially. I will say that I have noticed, myself and a lot of very bright people have told Tom that they have found the problem in his analysis and they have made suggestions very similar to mine, and he has always called or emailed me the next day and said I reanalyzed it and the signal was stronger. So I think these are good suggestions. I do not think it is going to matter.

Next slide. If the analyses remains positive, I don't think it is ever going to be possible to differentiate increased health care seeking behavior among families whose children are vaccinated on time. I don't think we have the capability of doing this, and suspect that the same finding will be replicated in the Harvard Pilgrim. If the people at Harvard Pilgrim can do it by June 21st, I will be amazed.

I am more worried that Harvard Pilgrim won't have the power. That the signal will fluctuate up and down so much, that we really won't know what to do with the results from Harvard Pilgrim. I think that's an egg in a basket and I don't think we should wait for it.

Next slide, please. So this is what I am proposing. I actually think we should do a cohort study using the population that perhaps we have already. We have got to define the population to study based on their known vaccine history, so their known exposure. You don't have their blood levels at the time of vaccines, but we know their vaccine history and we could do it. We could select Northern California Kaiser, Group Health Cooperative, Northwest Kaiser and Southern California Kaiser, and then

we could measure their outcome using a carefully measured set of neurodevelopmental and neuropsychiatric tests at one or more ages. This is not nearly as easy as it sounds because what we are really concerned about is exposure at one month and three months. But we may also want to know how about at one month and not at three months? How about two months and not one month. It would be difficult to do that while preserving enough power to see an effect at all the different vaccines levels. Because I am talking about bringing in children who had zero Thimerosal levels, 37.5, 50, 67, 75 and maybe truncate it there. So we have five exposure levels and then we have a two by two design. So what we would need to think about is actually, this is probably 500,000 children. I think we could find enough. It actually becomes a matter of findings rather than our ability to find the children.

Let me go one more slide. I want to talk about why am I proposing this? I think this actually breaks the link that probably exists in the observational study.

Children in this proposed cohort could have seen health care providers many times or actually never. We don't care about them. Our analysis of the neuropsychiatric and neurodevelopmental outcomes is no longer dependent on the parent bringing them in. We are going to insist that they come on in. Hopefully we will have good participation, and we are going to study them at six years of age, regardless of whether they never saw a doctor at all. So we are actually breaking that link. So we going to give each child now an equal chance of having the outcome, aside from their Thimerosal exposure which is what we are studying in order to find their populations to study different exposures.

There may still be some confounding because people who went to some clinics may have gotten very little

Thimerosal levels based on certain characteristics of the clinic they attended, and that may be an observable confounding variable.

Other confounding information could be potentially collected at the time of examination, including socioeconomic status, pregnancy exposures, smoking, alcohol. Some underreporting and under ascertainment of confounding will certainly exist and will certainly be diluted over time, but I have no reason to suspect that there will be a differential dilution or under-ascertainment by the Thimerosal sets.

To answer Walt's long standing question, I doubt this will allow us to differentiate Thimerosal. A lot of people have the same question. I don't think this will allow us to differentiate antigen number or vaccine number from Thimerosal, but it will get us a lot further down the road.

We could draw blood, and actually I would encourage people to think about drawing blood to look for gene environment interaction studies, because there may be a set of children in here that are particularly prone to Thimerosal related outcomes.

Let's talk about the confounding that is slightly true. The early receipt of vaccine in this study, children who have high levels of Thimerosal now at one month and three months of age are likely to belong to parents who are different, but why are they different? They are different because they are much more attentive. They are much more on the ball. I am really struggling here to use the term that is politically correct. The only term I can think of is the smarter parents. So actually what is this going to do? This is actually the confounding that might exist, although I don't know. The confounding that will exist will be a negative confounding. This is the children with the high

Thimerosal at the late ages, who are likely to be perhaps from better parents. The neurodevelopmental and neuropsychiatric outcomes are likely to be better.

I am not suggesting that as a reason to do the study. I am just pointing out that if you think about where the confounding is going to be focused in this particular study, the one that I would be worried about, the validity of the study, is that it will have several outcomes. That early high Thimerosal exposure will be associated with neuropsychiatric scores.

Dr. Verstraeten: One thing you can easily add, one arm would be to compare DTP-HIB combined versus separate, and I think with very small numbers you will have enough power, doing this kind of testing, to identify the difference between mercury, Thimerosal and the other, because they have the same antigens, the same amount of aluminum and probably a lot of the other stuff that is in the vaccines.

Dr. Rhodes: As you are thinking here, I think it would be important to sample from both cases and not cases though where you think a trend stands. I would argue for taking samples of cases and not cases.

Dr. DeStefano: Well, this is a cohort study, so it kind of separated it completely from cases.

Dr. Rhodes: But I think it would be important to say in those who have been called cases, do your tests pick up anything?

Dr. DeStefano: Let me argue from just strictly a pragmatic point of view. We actually are not looking for cases, we are looking for minute differences in neuropsychiatric and neurodevelopmental outcomes. I think dichotomizing people into cases, while serves a very sophisticated sample

scheme, may actually not be what to do for this particular one.

Dr. Rodewald: I would be more cautious about your ability to pick which direction the confounding is going to go. I think, for example, the birth dose of HepB is not usually quite processed the same in the U.S. and that is something that might be the other way. Trying to weight the preponderancies, in assuming clinic policies in terms of how early they get kids in and perhaps physicians are more worried about certain parents getting them in there early for this. I just think it is really difficult.

Dr. Davis: Yes, you said earlier the adoptees of the HepB. HepB is what you are saying.

Dr. Rodewald: And even getting kids in early after they are born. I mean a lot of pediatricians get the kids they are most worried about in earlier. That would go the opposite.

Dr. Davis: Right.

Dr. Rodewald: And the pediatricians opinions and practices dominate over parents preferences in terms of vaccinations and we know that from studies.

Dr. Oakes: I am going to have to leave momentarily, but first, is it feasible to contact people in these cohorts and select samples based on either exposure outcome history and go back to them?

Dr. Davis: I've not done it, but everybody is saying yes.

Dr. Oakes: It is ethically and practical to do this? So I would certainly argue to doing some kind of master case control study. In that case maybe on small groups. I was wondering if you

would have any idea on the numbers you would need in your cohort study?

Dr. Davis: No. I really don't think you can get enough to do a control study. I really think we have to move away from the idea that we can actually ascertain cases here. I think what we are really looking for is an age of...

Dr. Oakes: The master case control would have a different purpose. It would be to try to get at the ascertainment bias and other confounders. I'm sorry, I was putting two things together. That is a different issue, but I think if it is feasible to do that, even on a fairly small number of subjects and we want to do that, the people would have the hardest possible outcomes and some carefully thought out matched sample of controls.

Dr. DeStefano: The problem you would have is like the cases, we are still going to have to identify not only the cases we know about, plus we screen. You'd still have this problem of ascertainment that have been identified as cases now in our database.

Dr. Oakes: Well, you would know if they were really cases after...

Dr. Weil: Let me just before you get too far. You are going to run into some big ethical problems if you try and identify people from this study for some characteristic. If you look randomly into the study, the ethical problem won't be great, but the Human News Committee, if it is any good at all, is going to give you a very hard time if you try and identify people who are by number code only because they have a finding. That will violate all the rights to privacy.

Dr. Davis: In other words we are actually identifying them based on the vaccine.

Dr. Oakes: So it wouldn't be feasible to do it based on an outcome?

Dr. DeStefano: It wouldn't be, but I think we are still concerned about the ascertainment bias still being...

Dr. Davis: Recreate this, the bias that I am trying to get rid of.

Dr. Johnson: But isn't that what you are focusing on here? You are going to take what is called the positive endpoint and you are going to see if it is real.

Dr. Oakes: Because you would sample the controls randomly and get them from those eligible. There be no ascertainment bias.

Dr. Rhodes: I think the think is forget the automated outcome data. Go to the cohort and start with the exposure groups and the outcome can be defined upon the results of the tests.

Dr. Davis: So we are going to hire people to do careful neuropsychiatric and neurodevelopmental?

Dr. Rhodes: And I think at the same time you could, as an arm of this study, have people who are called cases in the automated data, people who are not called cases, and see whether these tests have any difference in those groups or not. Now, that would be interesting to know.

Dr. Davis: And you would be checking your analyses, but I would strongly urge you not to guess.

Dr. Verstraeten: To answer your question on sample sizes, depending on the type of tests and the difference you want to detect, sample sizes range from like 300 to 1,000 only. I think that is pretty close to the sample size in the Seychelles.

Dr. Davis: Is that for the entire study?

Dr. Verstraeten: The entire study.

Dr. Sinks: A couple of comments. I think your proposal goes very well in line with the beautiful studies that have been done, both in the Seychelles and the Faeroes, which are exactly this type of thing where you start from the exposure. You are not trying to determine case, you are trying to determine some difference in the neuropsychiatric tests, which has exquisite sensitivity and probably much greater power than trying to deal with cases. If you are going to do a master case control study, what I would recommend if you are going to look at cases is that you put a lot of emphasis into some standardized battery for determining who is and who is not a case, because I think that is one of the limitations with the data set that we have.

And I think the main purpose to do that particular study is to be more confirmatory of testing what Tom has presented to us. I think you are testing a very different thing which is, is this similar to methylmercury exposure? What we believe is there in terms of that biological plausibility.

Dr. Cordero: I would like to sort of follow up on what Tom said. It seems the question we really are asking here is does mercury or Thimerosal in vaccines pose a risk for selected neurobehavioral problems and therefore, I think having an accurate measure of ethylmercury is essential in whatever study. And just answering the question of having exposure by vaccines may be sufficient, the question still remains. Is it mercury or is it something else? So I think we are talking about measuring mercury and perhaps measuring other vaccine tendency exposures, then having some systematic way for looking at the outcomes and being able to classify appropriately what happened.

Dr. Stehr-Green: I just want to add one endorsement for a master case control study. In a case control study you have more

freedom to look at other exposures, more than just the Thimerosal in the vaccines. So it might be useful to do that. It might give you an opportunity to look at other plausible etiologies.

Dr. Davis: We are going to bring the kids in. Once you do that, your study cost has been incurred. If you do a one hour interview of diet and maintenance. So I have to do that anyway. Not to see the entire cohort.

Dr. Johnson: I think this has been an exceptionally useful and strong discussion, and because people have to leave we are going to have to cut it at this time. If we have time to come back before noon, we will.

I think there was a lot of recognition, and certainly I believe in most peoples minds, the implications of dealing with the composition of vaccines for the international community, and John Clements would like to make some comments at this time, then we will have Paul give us his rapporteur's comments. Then depending on the time we will come back to the discussion of research approaches.

Dr. Clements: Thank you, Mr. Chairman, I will stand so you can see me.

First of all I want to thank the organizers for allowing me to sit quietly at the back. It has been a great privilege to listen to the debate and to hear everybody work through with enormous detail, and I want to congratulate, as others have done, the work that has been done by the team.

Then comes the but. I am really concerned that we have taken off like a boat going down one arm of the mangrove swamp at high speed, when in fact there was no enough discussion really early on about which way the boat should go at all. And I really want to risk offending everyone in the room by saying that perhaps this study should not have

been done at all, because the outcome of it could have, to some extent, been predicted and we have all reached this point now where we are left hanging, even though I hear the majority of the consultants say to the Board that they are not convinced there is a causality direct link between Thimerosal and various neurological outcomes.

I know how we handle it from here is extremely problematic. The ACIP is going to depend on comments from this group in order to move forward into policy, and I have been advised that whatever I say should not move into the policy area because that is not the point of this meeting. But nonetheless, we know from many experiences in history that the pure scientist has done research because of pure science. But that pure science has resulted in splitting the atom or some other process which is completely beyond the power of the scientists who did the research to control it. And what we have here is people who have, for every best reason in the world, pursued a direction of research. But there is now the point at which the research results have to be handled, and even if this committee decides that there is no association and that information gets out, the work has been done and through freedom of information that will be taken by others and will be used in other ways beyond the control of this group. And I am very concerned about that as I suspect it is already too late to do anything regardless of any professional body and what they say.

My mandate as I sit here in this group is to make sure at the end of the day that 100,000,000 are immunized with DTP, Hepatitis B and if possible Hib, this year, next year and for many years to come, and that will have to be with Thimerosal containing vaccines unless a miracle occurs and an alternative is found quickly and is tried and found to be safe.

So I leave you with the challenge that I am very concerned that this has gotten this far, and that having got this far, how you present in a concerted voice the information to the ACIP in a way they will be able to handle it and not get exposed to the traps which are out there in public relations. My message would be that any other study, and I like the study that has just been described here very much. I think it makes a lot of sense, but it has to be thought through. What are the potential outcomes and how will you handle it? How will it be presented to a public and a media that is hungry for selecting the information they want to use for whatever means they have in store for them?

I thank you for that moment to speak, Mr. Chairman, and I am sorry if I have offended you. I have the deepest respect for the work that has been done and the deepest respect for the analysis that has been done, but I wonder how on earth you are going to handle it from here.

Dr. Brent:

Mr. Chairman, I think that was eloquent statement. The question that I have with regard to perceiving this data with some type of reanalysis, is that because of the diverse use on vaccination, no matter what you come up with somebody on one side will accuse you of doing something to get a negative result. Then if you come up with a positive result using the same data, the person on the other side will say see, we were right, it is causal. So I really encourage the investigators to get other populations to study because of the fact that I do not think reanalysis of this data is going to be as helpful as we would hope. It would be helpful if it wasn't in this room, because we know of the integrity of the scientists and we know they are pursuing it for the truth, but other people out there don't have those feelings about anybody who is involved in these studies. That is my concern and that is why I think Dr. Clements comments are so to the point.

Dr. Johnson: This focus on new research that has been mentioned that Dr. Clements' comments raised is the need, and this applies to the vaccine manufacturers to develop another, an alternative preservative anti-microbial measure for use in childhood vaccines than ethylmercury. It is possible in single dose. There is a lot of wonderful advances in manufacturing biologicals and it should be applied here I think.

Paul Stehr-Green, do you want to give us your rapporteur's summary of everything?

Dr. Stehr-Green: Let me say, my understanding is that whatever I say will be expanded upon once I have the benefit of seeing all the speakers notes and that a written summary will be submitted to at least all the consultants, is that correct?

Dr. Bernier: Yes, we haven't asked to do that yet, but Paul will be writing a report. We have a very short turn around for this. We want to get the report prior to the ACIP meeting, so we are looking at about a week to get this report. We want to get as much feedback from the eleven as possible, so if you could please collaborate with us in trying to get a quick turn around on that. So we would not want to get the report if you did not think it was a fair assessment, so Paul is going to have something as soon as possible.

Dr. Brent: But you want these written?

Dr. Bernier: Yes, I want you all to turn in your sheets, please. And also I would like to invite anyone else who has been in the meeting and heard this, I think the sheets have been widely available and anyone who has filled them out, please, we would love to collect those as well, even though we are focused on the eleven that were officially hired to be CDC consultants.

Dr. Brent: Who do I inform that my e-mail address is wrong?

Dr. Bernier: You can tell me.

Dr. Stehr-Green: Anyway, my point of reasoning was if you feel I have given an inappropriate slant or misrepresented comments you've made or others have made, we will have at least two opportunities to correct that, in this discussion and then when the written report comes out.

For the sake of time, when I write a written report my intent will be to summarize the sort of historical events that led up to this meeting. Both what has happened over the last several years and more specifically, Dr. Myers summary of the workshop last August at NIH.

Of note, I think it was important, or at least I glean from Dr. Myers' presentation, that in fact the group last year made a similar recommendation to what John Clements just said, and that is you may not want to do this study because the results are not likely to be useful for resolving this issue and in fact may raise concerns and havoc in locations with which we cannot deal based on this study. Is that a correct interpretation? So I think it is important that he verifies what John said, and he provides the setting for when this study was embarked upon.

I also intent to summarize sort of the generic aspects of the Vaccine Datalink at CDC and how the operation is set up. But of course most of the emphasis will be on the Phase I study. What I hope to do is demonstrate that through exhaustive analyses and very careful attempts to tease out a variety of problems with confounding possibilities, with other possible exposures, with other plausibles that we don't understand, with perhaps uncertain and inconsistent diagnoses and with this, to my mind, the looming issue of

the potential differential utilization of health care and the ascertainment bias that might carry with that.

Despite all those things, Tom and colleagues were able to demonstrate that there was a signal present, and I think the group verified that indeed there was a signal. However, that signal was not strong enough either by itself and in the context of others such as biological plausibility and so forth, it was not strong enough to support an inference of causal relationship. In fact it was a signal that deserved further investigation and that raised some perhaps disquieting possibilities.

I think in many respects the group of consultants has made my job a lot easier in that there was very little controversy in the conclusions. As a whole, the group was pretty unanimous, in fact we were unanimous, in saying that additional research is needed. However, that the current results were weak for a variety of reasons. Again, the inconsistent and uncertain diagnoses, the looming possibility of ascertainment bias and uncertainties as to whether or not we can separate out Thimerosal effects with other vaccine components, or even other exposures that may be somehow statistically correlated with vaccine administration.

Nonetheless, there was a consistent opinion that these weak findings should be followed up, and in fact in this last discussion we talked about different research avenues that might be pursued to get a better handle on this association, this signal if you will.

Again, with regards to the question of whether or not these results support causality, as I said before I think the group was unanimous, except for possibly Dr. Weil, in suggesting that there was not anything close to sufficient evidence to support a finding of a causal relationship. And again, we

went back to these issues of uncertainty about the diagnoses, uncertainty about the possible biases and confounding that could not be accounted for in the analysis because we did not have the data.

There were also concerns brought forward from previous human studies and animal studies that suggests the biological plausibility of this association may not be strong and supportive in that the calculated exposures in this setting with which this study dealt were actually below, in some cases quite far below, no effect levels that had been seen in human and animal studies previously, with a presumably more toxic form of organic mercury. So the fact that we were extrapolating from methylmercury to ethylmercury, the fact that we were extrapolating down a curve into an area where there had not been any observations of any effects, and yet still suggesting that there was this statistical association, it was my interpretation and it seemed the interpretation of others, is that the evidence for biological plausibility of this association was not very strong.

So in fact in summary, I think the mean list of the group was 1.8 in the rating scale. So as a group we said there is no evidence for causality for again the same reasons and recurring theme that came up.

In terms of the next steps, I don't have the same feeling of unanimity. I think this is a work in progress. I am not sure how we are going to resolve that, but we had some very good ideas put forward. The cohort study that Bob described seemed to have a resonance not only among the members of the panel, but also the wider audience here today. I think though that based on Bob's discussion and some of the comments that were made, there are a lot of issues that have to be resolved. How do we define exposure? How do we define diagnosis or the outcome

measures? How do we choose the subjects? How do we try and collect the information and or if we can collect the information, control some of these biases that have troubled these discussions over the past two days?

I guess my sense of it was that in terms of what the next steps are, we got some good ideas that have been put out, but they are pretty rough at this point and need further refining, and maybe that is not the role of this committee or these consultants, but I get the feeling as though we have resolved many of those issues. We had some promising avenues and the doors were starting to come open, but we have to peek through those doors or maybe walk through those doors and begin to feel some issue that are going to arise.

So I will leave it at that and see if people feel I hit the nail on the head or missed the mark or would like to add or subtract.

Dr. Weil:

Toward the end you made the statement of the 1.8 level or whatever it was indicated no causal relationship. I don't think that is quite true. I think it indicates there is no agreement that there is a significant sense of a relationship, but it might say the people felt there was some relationship somewhere. There is something in these data that relate the number of micrograms of mercury, or at least the group that is represented, that group seems to be related to something called speech delay. The data was significant.

Dr. Brent:

Causal.

Dr. Stehr-Green:

What I hope to do is draw distinction between findings of association and findings of causality, and I think the group, based on the answer to the first question, said that there was a finding of association. Perhaps weak, but there was a finding of association that needed to be pursued, but then

when you considered the issue of causality, the group was not willing to say there was sufficiently strong evidence to support a finding of causality.

Dr. Weil:

Or refute.

Dr. Stehr-Green:

Or refute, right. When I restate it that way, is that a fairer, truer characterization?

Dr. Rapin:

There is also the question of relevance. I mean is this tiny change relevant clinically? This business, you raised the IQ point by one point over a large population, it is statistically significant, but is it relevant? Can we measure the IQ that accurately, that this one little point is relevant? I think that is another matter altogether.

Dr. Weil:

Now they are reducing lead from 10 to 5, that is exactly the argument that is being used. That reducing acceptable lead levels from 10 to 5. The point is that is being discussed as a real possibility and it is based on a very tiny increment.

Dr. Rapin:

I think the whole lead issue to be revisited.

Dr. Weil:

But there is in other words another toxic compound that need to be looked at for some of these same reasons.

Dr. Rapin:

Even in my grandchildren, one IQ point I am going to fight about.

Dr. Johnson:

Paul, the hardest job anyone has at a conference is to be the rapporteur, and I am impressed. You are on top as far as an overview of what went on and you will get more of the written pieces. Yes, Bob?

Dr. Chen:

Before we all leave, someone raised a very good process question that all of us as a group needs to address and that is this information of all the copies we have received and

are taking back home to your institutions, to what extent should people feel free to make copies to distribute to others in their organization? We have been privileged so far that given the sensitivity of information, we have been able to manage to keep it out of, let's say, less responsible hands, yet the nature of kind of proliferation and Xerox machines being what they are, the risk of that changes. So I guess as a group perhaps, and Roger, you may have thought about that?

Dr. Bernier:

We have not specifically thought. I would take this opportunity to remind everyone that we have now been working with this information for several weeks. I think the fact that we were able to hold this meeting the last two days is a direct result of the fact that this information has been held fairly tightly. I think it has been a privilege to have this meeting and we have other meetings like this. As difficult as the science is, there are two other equally tricky, complex challenges. The policy crafting has to take into consideration some very diverse and complex issues. There is another group that will deal with that, and then we have the communication and how we handle this, which I think I am no expert at, but seems equally daunting to me as the scientific and the policy issue.

I don't think we can set a rule here because some people have gotten these documents. For example, some of the manufacturers were privileged to receive this information. It has been important for them to share it within the company with the experts there, so they can review it. Some of you may have questions. You may have given a copy, but I think if we will all just consider this embargoed information, if I can use that term, and very highly protected information, I think that was the best I can offer. If anyone else wants to make a suggestion, but I would say consider it embargoed and protected until it is made public on June 21 and 22 at the ACIP. There is a plan to do that.

There are policy groups that will be meeting before this, and communications experts that are meeting in advance, but until June 21 or 22, I think that would be the best way to proceed.

Now that I have the floor, if there is no other comments, on behalf of CDC I will take the prerogative of where I am sitting to thank everyone on behalf of the Centers for Disease Control, and probably on behalf of the Public Health Service, on behalf of the National Immunization Program.

If I could get to see some of you personally, I feel I made a connection with you when I invited you, but I feel bad that I haven't really continued that. I am looking at Dr. Rapin in particular. Dr. Stein. Some of you I haven't really had much opportunity to make contact with once you got here, but believe me, I am very grateful for what you did to make yourselves available, and I want to say thank you if I did not get to talk with you personally.

It may have been a blessing in disguise that the Super Comp Computer Conference was held simultaneously because it forced us to come to Simpsonwood, probably the only place in Atlanta that had any room. I think it created a spirit in this meeting that I think we benefited from. The kind of informality and effort to really try to figure out, which to me was the biggest challenge, what is the best way to understand and think about these observations. We really didn't know that when we came in here. I think we made progress as a group. That we have a better idea about the best way to understand these data.

The other thing that I was struck by was the quality of the science. Many of you commented very positively about the work that was done by the scientists. I am a proud member of the National Immunization Program. Prouder after this

meeting than I was coming in, and I want to congratulate the team at the table. I think you made us all proud.

The other thing I was struck by was the aura of seriousness, an implication that sort of hovered over all of this. Although we were all informal and this place gave you a feeling of a special spirit, I think overall there was this aura that we were engaged in something as important as anything else we have ever done. So I think that was another element to this that made this a special meeting.

I also think it has been extremely productive. Despite some of the semantic differences and issues that arose, I think in the end we stopped talking about that pretty quickly, and whoever suggested let's just keep going probably made the best suggestion. I think the questions in the end worked and I think we have had a productive meeting and that when we look at your notes, we will find there is a lot there.

I wanted to end by mentioning about the policy work and the communication work. I have also been struck by how much that is going to be as challenging. I have already said it, so I won't dwell on that, but you get the point that this is only one leg of a three legged stool and there are two other meetings just like this one that should take place, on the policy side and the communication side, as those experts try to get it right from their perspective.

Dixie, Walt, do you want to add anything?

Dr. Snider:

Just briefly let me say first of all, thank you very much. It has been interesting for me over the past seven or eight years in this position to go through folic acid fortification and Rotavirus and all the other interesting issues. I think this is one of the tougher ones. The fact is that your consultations make a tough job a bit easier and we are most

grateful to you being willing to come and contribute and we appreciate it greatly.

Dr. Orenstein:

If I can just add my thanks. One to the NIP staff who have worked and labored day and night for months to come to the presentations. I would like to thank Bob Chen, Frank DeStefano, Phil Rhodes and especially Tom Verstraeten. I have seen him in audience after audience deal with exceedingly skeptical individuals and deal with them in a very calm way in answering their questions and doing the analyses and I think you are mature well beyond your years.

I would also like to thank Roger Bernier who pulled off this meeting in rather short notice, and I think as everyone has said, I think this was an excellent meeting and is going to be very, very helpful to us, and we appreciate the time and effort you have spent.

In a sense this meeting addresses some of the concerns we had last summer when we were trying to make policy in the absence of a careful scientific review. I think this time we have gotten it straight. We've got the scientific review, because the policy and communications really have to derive from that scientific review. We appreciate all that you have done to help us with that and I think we will take it forward in working with the ACIP and other groups and agencies to try and carry on.

I would also like to thank Dick Johnston and Paul Stehr-Green for being the rapporteur.

Dr. Johnson:

Thank you, Walt.