what the detection limit should be based on our

years of experience in dealing with consumer

reactions and things like that.

We set a certain level with the kits we

developed; we picked 2.5. It seems to have gone

very well over the last seven or eight years since

these kits have been on the market. Some of the

other companies have a little bit lower range of

detection limit, and that seems to work okay, too.

However, if you go way too late, I mean,

they can all push these kits really, really, really

low. The problem is, Is there clinical relevance

at that point?

If there is no clinical relevance,

companies may be chasing molecules around their

processing plant. They will have all of this

positive data at a low level, and they won't know

what it means. We like to call this "paralysis by

analysis."

We want the data to be relevant. We want

the data to be useful. If the industry goes back

in and says, "I want to fix this, but what if I get

all of these positive results at a low level?"
Detection limits have to be kept in mind. They
should be tied to threshold levels, whatever we
decide the threshold levels should be.

It adversely affects the quality of life
for food-allergic consumers, if you use detection
limits that are really low or push those detection
limits without a good clinical basis. Because of
the industry reaction in the form of increased use
of "may contain" label.

When they did paralysis by analysis and
they get positive results, maybe they throw a lot
of "may contain" labeling on that product that they
are worried about and so they are going to put that
on there. That decreases the number of foods that
allergic individuals can eat.

The current detection limits that are set
that the industry uses right now have worked very
well for seven years in protecting the
food-allergic consumer.

I don't think at this point there is any
need to change them right now. But, again, as

science comes in and we know more about threshold levels, there might be an adjustment here or there.

We just finished an egg threshold study. Contrary to what Robert Wood said, you can do threshold studies in kids, because we did this in 30 egg-allergic children. That is the only kind of people we could find to have egg allergy are kids.

When we crunch those numbers and look at that data, if the threshold is low enough that we need to adjust the kids for egg out there, the manufacturers have all said they would be willing to do that based on the science.

(Slide.)

DR. HEFLE: Many companies are testing for allergen residues. What they are primarily testing is not-finished product, but they are using it to verify sanitation procedures. They have been using them for as long as they have been on the market.

Certainly with the new law coming up, there are a lot more using them than used to use them. In general in the U.S., companies are incorporating testing using these test kits. As

the test kits get faster and easier to use, it is easier for them to use them.

Again, the ELISA or lateral flow, which is kind of like a dipstick method are the preferred methods. Some do the test in house. They really like it if they can do that because they can fix things right away.

However, if you don't have in-house capabilities, they will send it out to a contractor lab or if they want third-party verification, they will do that.

Most companies, as I said, are not testing finished product. They are testing to validate

sanitation methods or doing environmental swabbing to try to find where the problem is before they get to the final product. They want to fix the problem before they get there and figure out if their sanitation is accurate before they get to the final product.

Some testing of finished product on certain occasions though is done, especially when you can have the product under full control. They

don't usually want to release something that they
have tested and they find out there is a problem
and they have to call it back from the marketplace
later and perhaps put consumers at risk.

There are tests that are based on DNA
detection, and they are called "PCR." We don't
advocate these for allergenic residue detection
because it doesn't prove the presence of the
protein. You need the protein to have the allergic
reaction. It just says that there is DNA from that
particular allergenic food there.

It is not practical at all for in-plant
use. You can't put one of these machines next to a
processing line. It is very expensive and requires
a lot of segregation and things. It is meant more
for a regulatory agency or a big corporate lab who
has this ability. It does not prove the absence or
presence of the protein or the allergen. It is
just an indirect kind of a marker.

There are ATP tests out there. This is a
test that is commonly used in the industry for
sanitation assessment. Some companies would like

to use this to detect allergens, specifically the ATP does not detect protein also. Right now, it is not knowing that these correlate well with the more specific protein-based tests.

(Slide.)

DR. HEFLE: Three peanut, like ELISA test kits, have been performance tested by FDA through AOAC-RI. Those companies with those tests are Neogen, R-Biopharm, Tepnel.

Five peanut ELISA kits have been studied in one JRC interlab trial. This is the European Union's group in Belgium that does these sorts of things, and they put these three tests plus two more through a validation trial. They are currently doing another validation trial on the two peanut lateral flow devices. They are not finished with that yet, but they are only doing one matrix not several matrices. They are just testing it in cookies right now.

(Slide.)

DR. HEFLE: FDA works with AOAC and has said they plan more validation studies with other

test kits, and that has been the case for more than
a couple of years now with no apparent progress on
this front, though.

The U.S. food industry and other
regulatory agencies -- for example, the Canadian
regulatory agency, the JRC -- has moved way ahead
of FDA/AOAC at this point. The industry is not
running validation trials themselves, but they run
in-house validation things like that.

However, there are regulatory agencies who
have said, "Well, we're going to move ahead. We
can't wait for AOAC anymore. We have to get these
things done in validated interlab trials." There
are several trials that are planned right now
internationally to, hopefully, get some of these
things "validated" in the next few years.

The U.S. industry has been testing for
about seven years now, since the first peanut tests
came out. They have increased the amount of
testing each year and, I've got to say, have spent
millions of dollars once they've gotten test
results to change equipment, to make modifications,

for allergens specifically.

Before about 10 years ago, we didn't have any tests at all to do this. Since the tests have become implemented, they have used them to make changes in how they manufacture food.

Health Canada/CFIA has a Compendium of Food Allergen Methodologies. They crunch through a lot more of these kind of in-house validations that they do so that they can use them for their purposes. There is a Web site for that. They use both commercial and their own in-house methods.

(Slide.)

DR. HEFLE: Validation of kits, there are more JRC trials coming out of the EU more likely. We know of several that are planned, and other groups have them planned, too. Other groups are planning more interlab trials, some with kind of "modeled" foods.

A lot of these tests are done where you spike peanut into something else. It is not really like a manufactured product, so it doesn't really mimic the manufacturing process.

A "model food" is actually where the allergen is manufactured into the matrix, so that

it more appropriately represents what would happen in the food industry.

Therefore, those are king of challenging to make. You can't just make them in your back yard or in your home kitchen. You need to make it on an industrial level, so it can be quite an undertaking.

(Slide.)

DR. HEFLE: Kit companies do much more extensive validation than ever will be done by any regulatory agency or academic center. It is usually that the are in the process of selling kits, and they don't necessarily share the data like they should. I have been encouraging all of them to go ahead and publish all of this great data they have, and it would be a lot easier for all of us to evaluate how good their kits really are. So far, they still want to sell kits and not spend time writing papers.

However, they do have liability issues.

Their kits have to work. They have liability

issues. They have reputation issues if the kits

don't work, so it behooves them to do their own

validations before they put a product on the

market.

(Slide.)

DR. HEFLE: Reference materials are solely

lacking for allergens. It would be really nice if

we had a bunch of reference materials we could do

all of these interlab validations with.

However, we are having a problem finding

the appropriate reference materials. There are not

many available, and they are really needed. NIST

is one source of reference materials.

Unfortunately, the NIST standards that are

available were not made for allergen testing, were

not designed for that and often do not represent

the type of allergenic materials used in the food

industry.

A case in point was the standard that was

used in the AOAC-RI-FDA study. It was peanut

butter made by a major manufacturer. It is fine

for things like aphlatoxin determination and other
things. Unfortunately, the varieties are not known
with certainty, because the manufacturer wouldn't
tell FDA about every little peanut that might be in
there. They wouldn't divulge it. I'm referencing
NIST not FDA, I'm sorry.

Different peanut varieties have different
responses in the kits. It is imperative to know
exactly what is one of these standards.
Unfortunately, there aren't a whole lot of other
standards around the world around to do that.

There are other sources of materials that
could be used as reference materials, but we have
to come to a worldwide decision on what is the
appropriate criteria for considering something in
the reference material. Is something the JRC makes
in Europe representative of something we use in the
United States?

There are several of these materials
available, and we could begin to talk about going
through some interlab trials with some of these, if
they met certain criteria.

(Slide.)

DR. HEFLE: Processing can have a huge effect on extraction and kit performance. Most kits are not validated using these model foods, so we have to do some more of this stuff; international call for more use of modeled foods.

The old method of spiking, which is where you put a peanut extract into some of the matrix and mix it together and see how it performs. This, again, does not truly represent what happens in the food industry.

However, the spiking does provide some useful information, but the manufacturing of these model foods gives the best information about how a kit will work.

Model foods have to be made on a pilot, plant or industrial size scale. If you make this in your backyard or your kitchen, then it doesn't really appropriate what a model food is in the industry, either.

If you make many cookies in a home-size oven or a Suzy Homemaker or Easy-Bake Oven, it is

not going to be the same thing as what Keebler or what Pepperidge do on a huge scale.

The results of these are not practical or useful for the food industry. Let's make some real model foods. They are involving for assessing how a kit is going to work with a specific commodity, how efficient the extraction method is under industrial conditions.

It is becoming more and more important to use these types of standards in assessing the kit's performance for certain commodities and processing. I think spiking is pass.

I get yelled at in my professional society, AOAC, because spiking is the way of the food chemists. However, we have to do some spiking and look at things, but we have to make these model foods and do that sort of assessment also.

(Slide.)

DR. HEFLE: The extraction method, is it sufficient? We've got to think about it. Is it sufficient? Is the recovery good? Can we trust the results?

Some foods are challenging. There are tannins and polyphenols in dark chocolate that bind protein. It is a famous matrix, one of the most difficult matrices to do with allergen determination.

High fat levels can hide the allergen in other types of ingredients. If the product is hydrolyzed, you cannot analyze hydrolyzed or fermented ingredients in these test kits. They were never designed for this. When you start chopping up the proteins, the ELISA signals go away. The methods are meant to detect intact proteins and not peptides.

Processing, if you burn stuff, it is going to be less detectible; it is less soluble. That is a factor. Now, most companies don't burn their food, but sometimes they want to detect burned foods on band ovens or something they can't readily clean. These are challenges to kit performance, too.

(Slide.)

DR. HEFLE: Most kits for most allergens

have good reactivity with processed forms of the allergenic food in my experiences over the last 15 years, and that is just my experience.

The use of polyclonal antibodies and crude extracts and making antibodies against processed forms are recipes for successful kits. There are several on the market today that do very well.

Monoclonals are okay if they use a heat-resistant epitope in making the monoclonals. They can accommodate the processing changes that occur.

Some of the egg residue kits have some issues in this regard. The industry has been able to adjust and adapt. Many survey the raw material instead. Instead of worrying about the processed egg, they will just do the raw egg and handle it that way, or use a kit that has antibodies against raw and processed egg, to get around that particular issue.

Matrix effects, my lab has used all of the ELISA-based test kits available on the market in our own validations and tests. It is kind of my

hobby so I like to do this. The matrix effects are usually not a problem for most of the test kits out there, for the vast majority.

Kit companies have added extraction additives to their extraction buffers to assist. When it was recognized dark chocolate was a problem, they added some secret extraction additives to help you pull the protein out of dark chocolate easier.

Model foods, though, again are going to be of great use in assessing the true extraction performance of a kit. Again, I can't stress enough we need to make more of these.

In cross-reactivity issues, even though most methods do use polyclonal antibodies, which those of you who know something about polyclonal antibodies could say, "Boy, there could be a lot of cross-reactivity problems with them."

We really don't see this happening. The kit companies really couldn't sell any kits if their peanut kit cross reacted with everything else, too. Therefore, we don't usually see these

problems in that they have looked at that before they have launched it, so we don't see the cross-reactivity. I am not saying that there isn't one that is going to crop up sometime.

(Slide.)

DR. HEFLE: Again, we've got a problem with hydrolyzed proteins, hydrolyzed vegetable proteins, hydrolyzed soy proteins. You can't really detect them.

The industry would love to do this, to chase them through the facility and see if they have cleaned up afterwards because we know there can be some residual allergenicity in hydrolyzed protein preparations.

However, the ELISAs are pretty much rendered useless when trying to analyze for hydrolyzed protein. It is not what they are designed to do. The company has had to make a decision, "What is most of our market?" It is not chasing hydrolyzed proteins, but it is chasing the intact proteins. We have to balance the kits to go towards that, so you can't use it for this.

Unfortunately, a negative result in an ELISA in this case does not mean that there is no allergenic residue left. You have to ascertain residual allergenicity via a different method using human allergic IgE in something like a Western blot or a RAST analysis.

Another related area is the analysis of fermented ingredients: gums, Lactobacillus cultures, starter cultures. Once they start eating at the substrate, the proteins are partially hydrolyzed and the ELISAs won't detect them anymore. You need to use an IgE-based method to just ascertain the true allergenicity.

Companies don't tell contract labs the nature of their samples. They just say, "Here is Sample X." They are not going to tell them it is hydrolyzed, so we have some challenges.

I try to communicate with the contract labs and say, "Be sure you ask the question. Just don't give them a negative result, because it couldn't be truly negative maybe from an allergenic standpoint." I think this is the minority of the

samples out there.

(Slide.)

DR. HEFLE: My lab performs testing for food-allergic consumers, their physicians, their lawyers when they call for free when they report a reaction to a food. We work with some members of the Food Allergy & Anaphylaxis Network when they have a problem.

If there is an analysis I can do, I will try to help a food-allergic consumer identify what happened with that particular food, if they have managed to keep it in the height of the moment.

In 10 years of doing this, we have only seen "large" -- now notice I say "large" with a quotation around it, I don't want to make a lot of judgments on that right now -- amounts of undeclared allergenic food causing reactions.

(Slide.)

DR. HEFLE: We cannot currently do immediate monitoring in the food industry, though. The technology doesn't exist. It is getting

better.   These lateral flow devices can sometimes

get down to 5 minutes now.   I think in the future

they will be able to make a more immediate

response.

Right now, a lot of them are 30 minutes

long.   If you are swabbing things and waiting

around for 30 minutes to see if the result is

positive and then having to go back and clean

again, it is pretty impractical for the food

industry to do.

Sanitation and verification is the most

practical, not the test and release kind of thing.

My dad is a fisherman, so I like to the catch and

release and test and release kind of analogy.

We do not have tests for some of the

allergens, and fish is a notable example.   You

cannot test for the hydrolyzed or the fermented

allergen sources using these types of methods.

Some types of cross-contact are not

homogenous or 100 percent cleaning is not possible

due to the nature of the product.   Food equipment

was never historically designed for allergen clean.

Sometimes these facilities are quite old, and there is no room. There is no room to bring in different equipment. They have to try to redesign as they can, but they can't get completely rid of hangup areas.

You cannot take enough samples to practically test, to be a hundred percent sure all of the time. That is impossible. If I get a statistician in to tell me how many samples I would need, the industry would just spend the whole day testing rather than trying to make food product.

In some of these cases, precautionary labeling is justified due to the nature of the product and the process in FARRP's opinion. For example, dark chocolate and milk chocolate on the same line is one example where we think precautionary labeling is justified. That doesn't mean we think precautionary labeling is justified in every case.

(Slide.)

DR. HEFLE: This is a study that we recently completed and published in 2004 of some

incidents from milk allergic consumer complaints.

These were the casein levels we found in those

particular products. They range from 5,000 on up

to 44,000 parts per million in things that were

supposed to be free of milk or labeled even

"dairy-free" or "kosher," quite high numbers of

parts per million.

They also asked me to talk a little bit

about highly refined oils. What does HRO mean? In

FARRP's opinion, "highly refined oil" means

neutralized, bleached and deodorized or refined

bleached and deodorized.

The definition of what "refined oil" is,

is kind of debated a lot right now in terms of

FALCPA, opinions based on scientific review of oil

challenges with oils in the literature and what we

feel refined oil should be.

The available quantitative methods, there

are methods used in the literature including ELISA

and other methods that reports the levels of

protein in highly refined oils. None of these,

though, have been validated in interlab trials or

other types of validation for protein and oil
determination to date.

Somebody will run something and they will
report it, and they will do a certain number of
samples, but no one has looked at whether that is
an appropriate method across the board for
detecting this.

There is a question as to whether a small
amount of protein in the HRO is completely
extracted in aqueous buffer. "Aqueous buffer" is
something that people often use to do these sorts
of biochemical tests. It means trying to partition
the proteins from the oil into an aqueous buffer.

If they really like oil, they might not
all come over. They might want to stay in the oil.
The question is, Does this capture the true protein
content of the oil or whether some of the more
hydrophobic proteins stay in the oil fraction, and,
therefore, do not get extracted and therefore
determined?

My lab uses an amino acid determination
based on Edman degradation, but we also use aqueous

extraction. We try to maximize that aqueous extraction.

We use heat; we use a large amount of buffer; and we concentrate the sample. However, I cannot guarantee that I'm pulling all of the protein out of that highly refined oil when I measure that.

We report the results as relative and not a complete picture of the possible protein count out of HR oil. I still think you are capturing most of the protein that is there, but I just can't sit up here and say we are covering a hundred percent of it.

(Slide.)

DR. HEFLE: The protein levels of HRO are reported in the literature, and there are lots of different reports and levels. The caveats again: The use of aqueous buffers in the determination, how good if they use an immuno-chemical-based method is the epitope recognition of the antibody? Does it really recognize those soy proteins at that level of processing?

Relating "total" nitrogen, sometimes they use the total nitrogen amount to what the protein is. Well, total nitrogen can be free and running around in the protein and not associated with -- free and running around in the oil and not associated with the protein. Consequently, it may be an overestimate actually of the protein amount.

Limitations of certain types of methods like dye binding. "Dye binding" is a method that will bind to certain proteins preferentially and not bind to others as well. When you use a dye-binding method, is it really representative of everything that is in there? You can't absolutely tell.

The protein levels reported in the literature are usually a few milligrams per kilogram, which are a few parts per million. You will see some widely ranging estimates, though, from different investigators. A lot of times I question their methods sometimes or their ability to reproduce that particular result.

I think that is the end of my

presentation, and I thank you very much for your
attention.

CHAIRMAN DURST: Thank you.

Committee, do you have any questions or
comments?

QUESTION-AND-ANSWER SESSION

DR. MALEKI: Soheila.

DR. HEFLE: Soheila.

DR. MALEKI: Yes, Soheila Maleki. I was
wondering, just based on your experience and you
have been around a lot of industry, if there is any
kind of correlation or if there are any standards
between what the companies use to label "may
contain" versus "contains"? Do they use the same
2.5 parts per million that the kits provide as a
may contain or a not contain and so forth?

DR. HEFLE: They don't really use the
analytical results to make a definite decision
about that. Usually, the companies make a decision
to put precautionary labeling on through a certain
stringent set of criteria. It is something they
have tried to clean up, and they are still having

issues.

They have intermittent contamination. They would never allow something that consistently had a significant amount of allergen in it to be called a "may contain." They would try to clean up more, if it is not supposed to be there.

They don't set a level like that. They use the analytical results to help them determine whether that is justified or not. It has to be potentially hazardous, intermittent, hard to clean. Those sorts of things are taken into consideration much more than just the simple analytical result.

DR. MALEKI: Thank you.

CHAIRMAN DURST: Yes.

DR. NELSON: Mark Nelson. I just wanted to follow up to that in response to Soheila. In 2001, the food industry, a group of associations representing their members did put together guidelines on labeling.

The preference is obviously and clearly the requirement is to label the ingredient in the presence of an allergen when it is directly added

to the food.   In the situation where there is a

potential for cross contact, we did establish some

guidelines before companies should use "may

contain" labeling because of the concerns we have

heard about before.

One of those key guidelines was to make

sure that we could not avoid it even after applying

good manufacturing practices: appropriate cleaning,

appropriate separation, and so on and so forth.

DR. MALEKI: I see.   Depending on how much

you detected, it didn't matter, if you detected, it

went to "may contain," if it was on the line or --

well, if it contains it was directly added to the

product?   I'm trying to make sure I understand that

correctly.

DR. NELSON: Yes, I think it is more to

Sue's point that we aren't necessarily measuring

the finished food so much.   It is not a catch and

release situation.

DR. MALEKI: I see.

DR. NELSON: It is understanding your

system; what ingredients are going into the food;

what other products might be made on that line;

validating your cleaning processes between

products; scheduling products, depending on the

ingredients that they contain; the sequence in

which you might make the product and so on. There

are a lot of things that go into it.

CHAIRMAN DURST: Anything else?

Yes.

DR. CALLERY: Pat Callery. It appears

that the allergens themselves are not that well

defined, especially when you can find in actuality

generated new allergens by treating food in a

certain way. I am wondering how you address the

analytical problem of false negatives and false

positives?

DR. HEFLE: For a lot of foods the

allergens are indeed known, and there are very rare

cases where you make new allergens through

processing. That is an extreme case in the

literature, I think.

However, false positives and false

negatives are evaluated at the company level first

by testing tens of thousands of food commodities
and looking for potential issues. Also, I kind of
poke around myself and see if there is anything
that I can challenge the kits with.

In my experience, the false positive/false
negative rate for most of these methods is very
low. I can't give you a number. I can't tell you
how good that is, because I haven't done a
systematic study.

However, I think that the use of these
interlab trials with model foods will help us look
at some of those issues a little bit more, but I
don't have a good sense of how much false positive
and negative is out there.

I just know in our experience, and we use
these every day, we don't have a lot of issues.
When the occasional issue crops up, and we call the
manufacturer. We usually work through it pretty
easily.

They do tell the manufacturers to validate
or run their own in-house validations before they
truly test the results. The manufacturers do tell

the manufacturers to do that, so, theoretically, they should hopefully find some of these things. However, every method has a chance of a false positive or a false negative.

DR. CALLERY: I'm not sure how you do that without standard materials.

DR. HEFLE: I'm sorry?

DR. CALLERY: I don't know how you do any of that without standard materials to validate them.

DR. HEFLE: Some of the manufacturers will give you a standard to work with, either the standards from the kit or a recognized standard or perhaps one of the NIST standards, which is what we are all defaulting to because we have nothing else.

DR. CALLERY: I think you mentioned that one kit, they have some secret materials that they put into the kit to help extract protein. This seems inconsistent with being able to validate a method if you don't even know what the test material, how it was made and what the scope of the antibodies are that are made.

DR. HEFLE: Well, the extraction additive is not a reference material. The extraction additive is just an aid in extraction. Usually, the companies will tell you what it is. It is usually non-fat dry milk or soy protein. It is secret, but it is not that secret.

It is just an additional protein in the mix that helps pull the proteins out of oily matrices or hard to extract matrices. The companies know this, and they share that with customers. However, these sorts of extraction additives aren't really the reference materials or the standards used in the kit.

CHAIRMAN DURST: Sue, will you be around for discussion this afternoon?

DR. HEFLE: Yes, I will.

CHAIRMAN DURST: I think we will hold further questions until that time because we are running a little bit late.

DR. HEFLE: Okay.

CHAIRMAN DURST: I would like to take the recess now. We will take a 10-minute break and

reconvene at 10:45.

Thank you.

(Thereupon, from 10:30 a.m. to 10:40 a.m., there was a pause in the proceedings.)

CHAIRMAN DURST: We will start with our next speaker, who is Dr. Stefano Luccioli, who is a senior medical advisor to CFSAN, FDA. He is also assistant professor at Georgetown University. He will be speaking on "Oral Challenge Studies: Purpose, Design and Evaluation."

ORAL CHALLENGE STUDIES:

PURPOSE, DESIGN AND EVALUATION

DR. LUCCIOLI: Thank you, Dr. Durst.

Good morning. Today, I really want to not talk to you as an FDA medical officer, but as an allergist who has experience in performing and evaluating oral challenge studies.

(Slide.)

DR. LUCCIOLI: The goals of my talk today are basically just to give you a basic overview of oral challenge studies, the purpose, why they are done, the design and conduct, and also spend a

little time on evaluation and interpretation of data, especially with regard to sensitivity of subjects as well as clinical response and severity and maybe present some data gaps that may be of interest while you deliberate on thresholds.

(Slide.)

DR. LUCCIOLI: The purpose of challenge studies are manifold, but the primary reason is to diagnose allergy, food allergy. The gold standard, as we have already heard, is the double-blind, placebo-controlled, food challenge.

As we have heard, also people outgrow their allergies. They are done also to evaluate tolerance where those individuals have outgrown their allergies. They have also been done to evaluate specific ingredients that are allergens in specific populations. For instance, there have been some studies on highly refined oils in peanut-allergic populations.

However, in recent years, there has been a lot of emphasis on using oral challenge studies to determine minimal eliciting doses. This has

important implications potentially to determine

sensitivities of individuals within a population,

but also potentially some therapeutic opportunities

in that, as Dr. Wood had mentioned, there is a

feeling that maybe if we can't cure food allergy,

maybe we can raise people's sensitivity levels so

that they may not react to very low trace amounts

of food.

For reasons that you are all here today,

also for establishing threshold challenges, they

may be able to provide you data on low-effect

levels and no-effect levels.

A problem in this field is that there are

insufficient animal models which are commonly used

to evaluate toxicologic ingredients and also

scattered data about case reports where there is

not a lot of information about exact doses that

cause reactions.

Very few studies are done or have been

done. One study was reported by Dr. Wood on

evaluating reaction severity, and we don't have any

current biomarkers to predict severity. This is an

important, I think, factor when we are looking at evaluating minimal eliciting doses.

(Slide.)

DR. LUCCIOLI: I'm just presenting this slide, but I'm not really going to go into it, to just give you an overview that oral challenge studies are somewhat different to traditional tox models that are used to determine potential thresholds or acceptable doses. I will, hopefully, be able to highlight some of these issues in my talk and present, as I said, some data gaps.

(Slide.)

DR. LUCCIOLI: When you are designing oral challenge studies, obviously the selection of subjects is an important factor. Usually, you have populations of adults, children or infants just to keep the statistics in order. Most studies involve both men and women as well as are from foreign countries and most high ethnicities.

The selection of subjects is basically geared to what the purpose of the study is for, whether you want to diagnose individuals with an

equivocal IgE or clinical history; evaluate
evidence of outgrowth of tolerance, as we have
mentioned; and also potentially to evaluate
co-existent allergies, for instance, milk-allergic
individuals who may have soy, especially in the
infant population and also for evaluating specific
ingredients, in this case how to evaluate infant
formula.

Obviously, for specific ingredients, you
may want to pick particular populations for that.
In fact, most infant studies are done to evaluate
infant formulas, and the majority of studies are in
adults.

Another important factor is that there is
a notable exclusion of individuals from these
studies. As Dr. Wood had alluded to, there are
individuals who have a cutoff level of their IgE
where above this level they have a 95 percent or
more risk of already failing the challenge. The
challenge is basically useless. You already have
the information, and you tell those individuals to
avoid the food.

However, these individuals may represent a fairly sensitive population. Now with IRBs as they currently stand, it is very difficult to get these individuals tested in studies.

Also, classically individuals who have had anaphylaxis or very severe reactions which were fairly convincing for the actual food are excluded from the studies, because another rule of thumb is do no harm.

Consequently, you don't really want to test people who could have potentially severe reactions when you have already had a high clinical index that they are allergic.

Of course, there are a lot of people who self-exclude themselves from studies who may be part of a sensitive population. I also mentioned here unstable asthma because in any study you don't want to test individuals who are unstable to begin with.

Individuals with asthma tend to have more severe reactions and are probably the group most representative of fatal reactions. By not

including these individuals, you may be missing not
only sensitive individuals but individuals who are
potentially very severe responders.

(Slide.)

DR. LUCCIOLI: With regards to test
materials, there is a variety of test materials
that can be used. Various preparations, if you
just look at peanut, you can have peanut flour,
ground peanut, peanut butter.

There is evidence that the processing
method of these various preparations may affect the
allergenicity profile of proteins within these
foods.

You may have some individuals who are more
sensitive to peanut flour versus peanut butter.
The importance, too, with choosing the material is
that for logistic purposes you want to have it for
an increased time, if you are going to be doing
challenges over multiple months or time points.

A preferred method for these types of
ingredients are dried ingredients. You get into a
problem where dried milk or spray-dried egg are not

very commonly ingested ingredients in the

population. It is more common, I mean, the raw or

cooked egg or milk, liquid milk. Therefore, these

are factors that need to be assessed.

Also, fresh versus processed foods, some

individuals are more likely to react to the fresh

food versus the processed as well as raw versus

cooked. These are issues that need to be

considered when you choose a food for a particular

challenge.

Then, the dose units are different within

these challenges. Some studies report milligram

for food; others milligram for protein of food;

and, very rarely, milligram per kilogram which

would be fairly ideal if we wanted to evaluate

potential differences between adults and smaller

adults, infants.

(Slide.)

DR. LUCCIOLI: Obviously, people who

partake in these studies are people who think they

have an allergy; may have had a fairly significant

reaction; and are, understandably, under a lot of

stress and are afraid.

Blinding is an important fact, since there
is unfortunately a high incidence of the "nocebo
effect," which is actually the opposite of placebo,
people reacting to a substance that they think is
going to harm them.

In blinding it is important to mask the
food, because you don't want the subject to know
what they are eating. Factors that are used are
called "vehicles" in one sense, and they are
basically other types of foods that are thick that
can hide the taste and smell and texture and that
are also pleasant tasting, you hope.

However, when you are thinking about doing
a challenge study over a few time limits, obviously
you don't want to give some of these vehicles too
much of this, too many milkshakes -- you have to
make sure that the individual is not milk allergic
-- but also they may cause some GI effects or other
things independent of what the actual food that you
are studying would have.

In some cases, they don't always mask the

taste. Therefore, some researchers have preferred to use capsules, since this basically bypasses the taste issue.

However, using a capsule is difficult, especially if you are going in higher doses of food, it is hard to put a serving of some food into a capsule. I think people would know when they see a big capsule that there is more food in that.

Also, an important factor is that you may delay the absorption of that food putting it in a capsule, and also you bypass the oral cavity which may be a primary target organ for the initial allergic response. You may have not only a delayed response but potentially a less severe response.

I won't talk about the protocol, I think that was basically well-mentioned by Dr. Wood, but also a question about placebos. There are some studies that use placebos within the challenge. They use a dose and then the next is a placebo.

You know, it is a very complicated process where you usually need some other people that blind those to both the researcher and the subject, but

they are used as well. However, I think the

preferred method and the easier method is to do a

separate placebo day.

(Slide.)

DR. LUCCIOLI: Now this is just a

schematic of an example of a dose protocol. I

think the important factor is this is an escalation

study of divided doses. One of the important

things, too, is you don't want to be there all day,

and you don't want the patient, too, to be there

all day.

To be able to determine a dose of food and

get up to the final dose, which is usually a

serving of the food, which is like 10 grams of

solid or 60 grams of wet food is what you want to

achieve.

If there is no response at that dose,

there is a good likelihood that the challenge is

negative. However, in many cases you still want to

have the patient come back and do an open

challenge.

Now, with choosing the starting dose, this

varies among studies. In many diagnostic studies, because of this issue about not wanting to be there, you choose a dose that is roughly half of the dose that caused the reaction.

Now, I don't know how a lot of people figure that out, but that is what has classically been used as the starting dose. Even within a study, these doses shifts. This dose usually comes out to be in the milligram range.

Now, more recent studies that have actually been targeted to study minimal eliciting doses, have started in doses in the even microgram range. However, there are a variety of studies when you are looking at evaluating studies for eliciting doses.

Also, in this protocol, it is important to know the time interval differences. Usually, also that is tailored to the patient when their symptoms first occurred. Most allergic reactions occur within 15 to 30 minutes, so that is usually the time gap, but some other reactions may be a little bit more delayed.

As Dr. Wood discussed, there are individuals who have delayed reactions as well. Unfortunately, it is just not logistical to do a study and wait for these people's reactions to occur, because they might not occur that day; they may occur on a separate day.

In this model that I use, I just use a twofold dose incrementation, but also this could vary. Some studies go up to even tenfold, so this could affect also the starting dose and interpretation of doses in the dose response.

Now, you go and you do the challenge. If it is negative, it is negative, or you stop it after the first objective symptom occurs. Some studies will also record the subjective symptoms, but that is not always the case, because the objective symptom is the symptom that denotes a positive allergic response.

When you record the dose, you can either record it as the 4X, which is the discrete dose recorded or the 7X, which would be the cumulative dose adding the X, 2X or 4X.

Just to put this also into some
perspective in terms of safety assessment, when we
are talking about LOAELs and NOAELs, the 4X would
be the low-effect level for this study. If there
are doses before that, at least for this individual
you can say that this dose did not cause a response
and could be considered a no-effect level.

(Slide.)

DR. LUCCIOLI: Some other issues are don't
do this at home. People can have a very severe
reaction. These studies are done in a clinic or an
office where there is emergency equipment and
personnel. It is not a challenge that is done out
in the open. It is in an experimental setting, so
that can also affect the interpretation or results.

Medications, too, most studies now have
people stop the medicines, but with some earlier
studies this was not a factor. Antihistamines and
other things, if people are on these drugs, may
block the early responses so that can factor in.

Fasting, too, most people fast before the
study, but in some studies this was not necessarily

explained. If you have a full meal right before

the challenge, this could affect, potentially

affect, absorption of the allergen and therefore

affect the interpretation of the study.

The clinical history or reactivity, too,

is important. Dr. Wood talked about oral allergy

syndrome, but he did not mention about exercise.

There are some individuals who eat a food and have

no problem. However, if they eat the food and

exercise, they have a problem.

Some studies actually test the individual

and then put them on a treadmill and have them

exercise to see if you can elicit the reaction. I

mean, this is very rare, but that is something that

also can be done in terms of the oral challenge

setting.

(Slide.)

DR. LUCCIOLI: Statistical endpoints, I

think these are fairly straightforward for most

challenge studies. You want to just know what

percentage of individuals will react or not react

to the challenge, or in cases where you are

studying reaction severity which ones will have a mild versus a severe reaction.

If you assume that all of these individuals in the study are part of the sensitive population or general population, you can maybe make some assumptions about that and decide a percentage that will or will not react to a specific food concentration.

Also, there an importance in this is also when you are designing a study, you may want to try to achieve a certain number of individuals to give you confidence levels for the incidence of allergic reactions.

In this example, this is a table that shows over here (pointing) the number of individuals that need to be tested to give you a confidence level that the incidence will be less than this.

For instance, if we were to design a study with 66 people, that would give us 99 percent confidence that 1 in 10 would potentially react, so 90 percent would not react. Also, you could use if

66 is more than 59, you could also say, well, 95 percent confidence that 95 percent, 1 in 20, will not react.

Twenty-nine has been usually seen as a magic number for infant formulas. If 29 patients do not react, if the infant with milk allergy does

not react to a cow's milk infant formula, that is a basis for hypoallergenicity.

(Slide.)

DR. LUCCIOLI: I will spend the rest of my talk on evaluation and interpretation of challenge

study data. Basically, a general interpretation as we just talked about the statistics, many of these studies are done in a very small population of patients, therefore you cannot make a very general assumption for the general population.

(Slide.)

DR. LUCCIOLI: Because some of these studies do test the same food, there is a tendency to group these studies together to try to get the power higher and then potentially make some assumptions.

The problem with this is that I think it
is important to note that all of the studies that
are currently available are not standardized. I
think that was a question asked just a little
earlier.

This is not standardized data. They are
not standardized to dose. Starting dose or
blinding or testing could also be a factor and also
interpretation of clinical symptoms, which I will
address a little later.

Another issue here is that all sensitive
populations, are they included. If you have
information only on adults, is that going to
predict what harm it will be to infants.

Again, in terms of statistical power, if
you get individuals who are not reactive, if you
are looking at total numbers to say "This is how
many people did not react to this dose," well, what
about people who didn't react to the challenge at
all? Should they be included in the final analysis
of individuals, or should only the ones who react
to the challenge be part of that analysis?

What about foreign study data. For instance, China has a very low prevalence of peanut allergy, presumably because peanuts there are boiled or fried versus in this country they are dry roasted. If you have all of this data in the United States about peanut allergy, could that be transferred to data in China?

(Slide.)

DR. LUCCIOLI: I just should mention, too, that with standardization it is important to note that there have been some very nice reviews on actually proposed protocols, standardized protocols, for food challenges which have been published in the last year or so. However, to my knowledge, there have been no studies that have used this protocol at least for a major food allergen for evaluation.

Another general interpretation is that this is an experimental exposure. It is not real life. There could be false negatives. Individuals who have had a negative food challenge go out and have an open challenge and react. It is not always

a definitive assessment of allergy. Also, I think
it is difficult to predict reactions to future
exposures. I will try to talk about that as we
come up.

(Slide.)

DR. LUCCIOLI: Subject sensitivity, this
is I think an important issue to consider when
looking at evaluating food ingredients. The
genetic heterogeneity of individuals, there are
multiple allergens in food.

People can be sensitized specifically to
certain allergens within that food. If you cook
the food in a certain way or process it, you may
affect their allergenicity positively or
negatively. This may be what is apparent when they
do studies and you see this enormous gap in
responders.

You have almost a millionfold gap between
the high responders or I should say the least
sensitive who respond to low doses and the most
sensitive to who respond to high doses.

There is also this potential link with

severity, as Dr. Wood study has suggested and some

others, that some studies suggest that the

individuals most sensitive to low doses appear to

have the most severe reactions. Are we talking

about a specific subpopulation of individuals here

who are not only sensitive but severe? Also, there

is a sensitivity issue between foods and between

food products.

Another important aspect is that the

individual sensitivities may vary over time.

Allergies can progress and individuals with food

allergies develop asthma later in life. This

asthma, therefore, makes their reactions a little

bit more severe.

Telling somebody right now that they

reacted at a certain dose and that it is okay to

ingest doses before that may not be relevant a year

or five years from now.

(Slide.)

DR. LUCCIOLI: This just is a hypothetical

dose curve adapted from Jonathan Hourihane, who has

done some nice research in this area, basically

just to show you how severity and sensitivity may

factor in. I don't really want to spend time on

that.

(Slide.)

DR. LUCCIOLI: Evaluation of clinical

responses, this is where interpretation of

eliciting doses is important with regards to

subjective versus objective symptoms as well as

reaction severity in the dose response.

This table summarizes some of the

reactions that you can see from an allergic

response. Basically, they are divided into

subjective versus objective. "Subjective" means

that they are reported by the individual or the

subject, and "objective" are responses that are

actually visible or observed by the observer.

These reactions are reported in this

manner. As I said, it is when objective symptoms

occur, that is when the study is felt to represent

a positive reaction and stopped.

(Slide.)

DR. LUCCIOLI: To just show you some of

these reactions, not only is there a wide range in
reactions, but there are some fairly milder
reactions, hives. You down here to shock and this
is anaphylaxis. Wheezing and syncope are very
close to systemic reaction and potential
anaphylaxis. Consequently, even within an
objective response, you may have a severe
anaphylactic response.

There are also some subjective reactions
that may be somewhat severe: throat tightness,
dizziness, sense of impending doom. I haven't had
the pleasure, fortunately, to experience a patient
with this, but I hear it is fairly dramatic. They
have this sense of impending doom and go rapidly
into anaphylaxis. It is very, very serious. It
doesn't take much for a subjective reaction to go
to something severe.

Also, there are some reactions that kind
of are in between the line of what is subjective,
what is objective: fussiness behavior, abdominal
pain. In adults, that could be suggestive of a
nocebo effect. However, in infants, infants don't

mess around.   This is their symptom, so these could

be positive responses for infants.

At the same time, you could have skin

flushing or shortness of breath leading to

increased respiratory rate, which could be an

objective sign.   However, many times this could be

due to also a nocebo effect.   Whether these are

actual positive reactions is hard to determine.

There is some clinical interpretation differences

that can occur here.

(Slide.)

DR. LUCCIOLI: Subjective versus objective

symptoms -- as I told you, the measurable indicator

of allergic response is the objective symptom.   It

has got many different endpoints, and the

interpretation may vary.   This could also be true

for the subjective reactions.

Many times, subjective reactions do occur

as part of a nocebo effect.   However, there are

some that are potentially indicative of an allergic

reaction.

How should these be factored into the

assessment? Many times they are not recorded in the study, so we don't know if there are earlier reactions to the objective dose, which may represent an earlier adverse event level.

(Slide.)

DR. LUCCIOLI: Some other eliciting dose considerations, the starting dose is important. If the response occurs at this dose, you cannot determine the no-effect level. Obviously, there is no dose below that that doesn't cause an effect, but is this starting dose the low-effect level? Could you have given a dose a little lower and they could have still reacted?

With dose increments, some are twofold and some are tenfold. Using tenfold, you may miss some increment in between that there could have been a reaction, even maybe a fivefold difference.

Also, time intervals between doses, as Dr. Wood has explained, some doses are delayed. However, time intervals, if you don't give enough time, you might not know when a subjective response has become a subjective response or so forth. This

could also affect interpretation of these eliciting
doses.

Of course, discrete versus cumulative
dose, some studies report just a discrete dose;
some the cumulative; some both, which is better.
However, how do these factor into a true exposure
assessment or prediction?

(Slide.)

DR. LUCCIOLI: I just want to just show
this, a few more slides, just to kind of put this
into perspective here, give you a mechanistic view
that allergy is a unique toxicologic response.

When you get food that gets challenged, it
causes a massive release of mediators and
cytokines. This is an amplification system that
the immune system uses to protect itself.

Now, in many cases, this response occurs
locally and may not amount to very much, but in
some cases this amplification can involve other
organs and spread systemically very rapidly.

(Slide.)

DR. LUCCIOLI: What has been observed is

that the severity of an allergic response is on a
continuum. You can have subjective responses at
some point, objective anaphylaxis, and potentially
death in worse cases.

A few points to note is that this is not a
fixed response. The early objective system may
rapidly progress to something worse. Also, the
degree of amplification, this is not always
predictable or reproducible, so symptoms may not
always be reproducible on subsequent rechallenge.

(Slide.)

DR. LUCCIOLI: To end, with the reaction
severity, most studies only report the actual
symptom. You don't know where this symptom is in
the continuum of severity many times. Those few
that do report the symptoms, they report them as
mild, moderate, and severe.

You have to interpret the researchers, I
guess, response to this, how they interpret it; so,
there is some interpretation. Also, when you have
severe response, like in Dr. Wood's study, in some
cases a third of individuals reacted and had mild

reactions, a third had moderate, and a third had severe. How do you factor in those severe responses when you determine uncertainty or other issues?

Also, potentiating mitigating factors are important: anxiety/stress, medications, and so forth. These can either potentiate the reaction or stop it.

Then, the challenge stops after the first response. A lot of times we don't have the luxury of knowing how far or how many more doses would have caused a more severe response. Having that information is important when you are wanting to make some risk assessment decisions. Again, it is a dose distribution, not a dose response.

(Slide.)

DR. LUCCIOLI: In conclusion, the oral food challenge does provide data on clinical sensitivity to minimal eliciting doses and also reaction severity to the initial dose. However, challenge data currently available for interpretation is not standardized among studies.

The current data pool may not include extremely sensitive populations with regards to severity. Challenges have a proven value as a diagnostic tool but less value in predicting reaction severity to future exposures.

Thank you, and I will be glad to answer some questions.

CHAIRMAN DURST: Thank you very much.

Are there any questions from the Committee?

We will start here.

QUESTION-AND-ANSWER SESSION

DR. BRITTAIN: Yes. I have a comment on the sample size table that you showed us. I am not sure that is incorporating the statistical power education we need to have more than these individuals. Are you familiar with what I'm talking about like the 29 there?

DR. LUCCIOLI: Yes.

DR. BRITTAIN: I'm wondering if you get zero out of 29, then your confidence interval excludes --

DR. LUCCIOLI: Well, what that 29 is, that
is usually a number that is targeted to challenge a
number of study subjects. If you show that 29
individuals with the specific allergy do not
respond to that ingredient, that gives you 95
percent confidence that 90 percent will not react.

DR. BRITTAIN: I guess what I'm saying is
that means if you observed 29, you get the desired
confidence interval. However, if you were planning
a study and you wanted statistical power to be a
certain amount, you would need to have a bigger
study.

DR. LUCCIOLI: Sure.

DR. BRITTAIN: You couldn't assume that
nobody would react.

DR. LUCCIOLI: Yes. Yes, I mean, you saw
that to be totally assured you would have to test
quite a few people.

DR. BRITTAIN: I do have another question.
You mentioned the placebos again, if someone does
have a reaction with a placebo, how is that
interpreted in terms of if they also have a

reaction?

DR. LUCCIOLI: Well, yes, many times you don't know, so then you unmask the study and then you find out that they reacted to the placebo. Now, technically, some studies will rechallenge that patient again. They will have them come back just to say, "Well, maybe" -- sometimes people do react.

The difficulty is when they react to the active dose, to a real challenge, and to the placebo. If the placebo is too close to the active, it may be that by the time you gave the placebo, they are still having the active reaction.

Basically, if they are rechallenged and show again, they are excluded from the analysis. Now, that is what should happen. Unfortunately, you never get that information a lot of times from these challenges.

CHAIRMAN DURST: Suzanne.

DR. TEUBER: One of the aspects that we are all very concerned about is which threshold to use and when it may cause a subjective reaction.

Actually, oral itching is a very important

subjective reaction that you didn't have on your

table up there in this presentation.

DR. LUCCIOLI: Okay.

DR. TEUBER: However, if that is

reproducible with two active challenges and not

seen with two placebos, which I think Dr. Taylor

may address a little bit later, but some of the

studies that Dr. Wensing and Bindslev-Jensen and

Dr. Hefle have been doing, they have been looking

at that. All of these have been followed by

objective reactions at higher doses.

DR. LUCCIOLI: Yes.

DR. TEUBER: I would really like people to

comment on that because this may be a much safer

way to approach obtaining thresholds to get these

extremely sensitive populations, if we can use

reproducible subjective data knowing, too, that

there are those other factors that may affect it.

DR. LUCCIOLI: Sure.

DR. TEUBER: For instance, in these

threshold studies that are being designed

specifically for thresholds, people with unstable

asthma would still be excluded. I am curious if

anybody knows anything about how unstable asthma

would affect the threshold for a LOAEL that is

seen? Is it a lawful difference? I mean, is there

any anecdotal experience with how that might

change? We want safety here.

DR. LUCCIOLI: Obviously, a speaker that

is coming after me would have some information on

that, but some information from Jonathan Hourihane

would suggest -- and I think some European studies

actually do test some severe patients. Now, I

don't think that any of these patients are

unstable.

I think that they are all excluding

patients who have unstable asthma, but with asthma

in general they haven't found that these

individuals have a lower minimal eliciting dose

than other individuals. However, when they do get

a reaction, they can have a much more severe

reaction.

The assumption, though, is that because of

the fatalities and other things that when their

asthma becomes unstable their sensitivity could

change and become more severe very quickly.

CHAIRMAN DURST: Marc.

DR. SILVERSTEIN: Marc Silverstein. I

have two comments that deal with sort of

clarification of terminology and one comment that I

think deals with a more difficult issue. I thought

this was a wonderfully helpful and concise summary

of a variety of complex factors.

In terms of the two clarifications, in

clinical medicine from the first days of medical

school we are taught the difference between

"symptoms," which are subjective, and "signs,"

which are objective.

Some of us from the clinical side who will

be reading the report will think that subjective

symptom is redundant and objective symptom is an

oxymoron.

To help the wide dissemination of the

report and presentation, I would like to suggest

that we in our thinking we may say "subjective

finding such as symptoms of the disease" and "objective signs" is the sense that I use that.

DR. LUCCIOLI: Sure.

DR. SILVERSTEIN: I think it is helpful because there will be a variety of readers of the report who may not appreciate that on the clinical side there is a clarification about that.

The second clarification had to do with the incidence versus prevalence in the sample size table. What we are talking about is the proportion of subjects being tested to the proportion of individuals in a population, so that sample size table or the table we have is the expected number of sample you would need. You label it "incidence" but it is really a "prevalence" of a condition in a population.

DR. LUCCIOLI: Okay.

DR. SILVERSTEIN: I believe that what you are getting at is the sample size so that the lower confidence interval is that 10 percent or 1 percent rather than the sample size necessary to show that two populations differ in proportion or the sample

size to show how tight you are around a rate of
zero, which would be a different population.

DR. BRITTAIN: Can I respond to that?

DR. SILVERSTEIN: Sure.

DR. BRITTAIN: I don't think when you are
designing a study you want to think of it in terms
of statistical power, which would be greater than
the sample sizes.

DR. SILVERSTEIN: The third comment I
have, which is substantive and I think we may need
to address this later in greater detail, has to do
with sources of error. There are two sort of
classes of errors that we made in our inferences.

One types of error an epidemiologist or a
clinical epidemiologist may say is biased, one of
the most common sorts of types of errors we could
make would be making inferences in the presence of
certain biases. The most common of which would be
selection bias.

Of course, the selection of individuals
who are referred to a physician, who are referred
to an allergist, who are selected for an oral

challenge, food challenge, study would potentially lead to erroneous inferences if there were non-representative selection.

That is something that in reading the literature and making decisions about inferences for studies or for policy I think we need to be aware of up front, so that is an important class of errors that we need to be alerted to.

The second would be an epidemiologist would talk about confounding. In your example of a study subject who has asthma, whether it is stable or unstable and how that is defined, asthma might be considered an extraneous factor that affects the relationship between the allergen and the response to the test. We could use the framework of thinking of it as a confounding benefit.

Factors such as bias and factors such as confounding, I think, are useful as we make decisions about the report and the evidence for that. I would like for us to be alert to that as we think about the presentations.

DR. LUCCIOLI: Thank you for the

clarification.

CHAIRMAN DURST: Are there any other
Committee comments?

(No verbal response.)

CHAIRMAN DURST: If not, thank you very
much.

Our next speaker is Dr. Rene Crevel,
senior scientist at Unilever, Safety and
Environmental Assurance Centre in the United
Kingdom. He will be speaking on the "Threshold
Modeling Approach."

THRESHOLD MODELING APPROACH

DR. CREVEL: Well, first of all, thank you
for inviting me to share some thoughts on the work
we have been doing on modeling thresholds.

(Slide.)

DR. CREVEL: You have asked me to talk
about the following, to look at different modeling
approaches including what is named the
"hyperallergen." This is the model, and the
Bindslev-Jensen, et al., allergen model; talk about
the data requirements and underlying assumptions

behind them; and then say something about

interpreting the results of applying these models.

(Slide.)

DR. CREVEL: Now, just to take a step back

and think about why we are doing this, we've got a

challenge in allergen risk management as far as

industry certainly is concerned.

We want to protect allergic consumers of

course, but we also are aware that protecting them

by certain measures of risk management such as we

have heard about this morning like precautionary

advisory labeling does actually affect their

quality of life.

We want to minimize the effects on their

quality of life, the adverse effects on their

quality of life. We ultimately also want to

maintain economic operation of food manufacturing,

because if that doesn't happen, then that will

affect the quality of life of a considerable number

of individuals and people throughout society as

well. It is an important point to bear in mind.

(Slide.)

DR. CREVEL: How can we meet the challenge? Well, first of all, of course we could label where the allergen is present, and that is fine. You have legislation over here now in the U.S. for that; we have legislation in Europe; and many other regions and nations also have legislation.

(Slide.)

DR. CREVEL: Or, we can ensure that the residual allergen content of product is low enough to be harmless. I put in brackets here (to the vast majority of allergic consumers), because we have heard here this morning some instances of people reacting to extremely low amounts.

I think it is questionable, and I think the Committee must address that particular question, whether those people can be protected by whatever we can do in the food industry. We need to think about what alternative measures may need to be done, whether they can ever eat the sort of foods we can produce.

(Slide.)

DR. CREVEL: How can we determine what is harmless to an allergic consumer? Well, we have several sources of data, which I have listed on this particular table.

We can look at case reports from the literature, and those we have heard. We have heard about people reacting to very low amounts. Unfortunately, the usefulness in risk assessment, actually an allergen risk assessment in my view is actually quite limited.

They do establish the hazard. Yes, they tell you that a certain amount will affect some individuals, will provoke a reaction in some individuals. They don't tell you in how many individuals that will happen.

We can use control challenge studies. In fact, those of course provide the bedrock of what is needed, the information needed in allergen risk assessment because the population can actually be quite accurately describe.

You can describe them in terms of the symptoms they have experienced, the allergological

history, or the medical history as appropriate --
all of the demographics that you can think about.

Finally, dose distribution modeling, which
I am going to spend obviously some time on, also is
very useful in allergen risk assessment. But of
course it relies on the experimental clinical data
which is generated in control challenge studies.
It cannot operate in a vacuum. We do not have
enough of those sorts of data at the moment.

(Slide.)

DR. CREVEL: I have been asked to say
something about the hypoallergenicity approach. As
I understand it, it is an unofficial standard for
designating infant formula as hypoallergenic.

The original reference I found goes back
to 1991, although I think the American Academy of
Pediatrics has actually updated or at least issued
the guidance more recently, I think, in 2000 or
something of that sort.

The statistics of this approach are based
on the binomial theorem, quite simply. This shows
that, for instance, if you have a study with 29

participants, as we have already heard, and you

observe no reactions, then you can be 95 percent

confident that only 90 percent of the population

from which these people have been taken will not

react.

You can also extend that a bit, so if you

observed one reaction and you added people to the

study, then you would 46 for the same degree of

confidence.

If you wanted 95 percent confidence then

fewer than 99 percent would not react, then you

have got those other numbers which already become

very challenging, pardon the pun, for a challenge

study both in terms of recruiting the people do it,

to participate, and the economic cost of actually

doing it.

(Slide.)

DR. CREVEL: However, those are very

useful data when they are generated, but protecting

90 or 95 percent or even 99 percent of the allergic

population is not sufficient as far as we are

concerned as an objective for the food industry.

What we asked ourselves is, How can we improve this? There are several ways. We could try to look at the conventional toxicological approach and apply a safety factor to the lowest observed adverse effect level or the no observed adverse effect level as the case may be, if you've got the no observed adverse effect level.

However, that particular safety factor, I would say, would be arbitrary because we don't actually know enough about interindividual differences within the allergic population to apply a science-based factor, I think. The level of protection still is undefined. You do not know how many people you are protecting by applying that particular safety factor.

What about modeling of those distribution of minimum eliciting doses? Well, that can actually define the level of protection for individual allergen level. You can actually use a safety factor there. You can use something which is like a lower 95 percent confidence interval instead of using the figure itself.

(Slide.)

DR. CREVEL: Does modeling actually work?

We asked ourselves could we fit a curve to the

distribution of minimum eliciting doses that are

generated by challenge studies, and could that

curve be useful to predict the number of reactions

likely to occur as a result of exposure to a

specified amount of inadvertently present allergen

in the food?

What I have to say, of course, is we are

not so much concerned about "declared allergen,"

people who are allergic can avoid that, but what

our concern is about is that which is present by

cross-contact, mainly inadvertently.

(Slide.)

DR. CREVEL: This just gives a very quick

model curve.   It is just used to illustrate some of

the points, right, okay.   From this particular

curve, okay, we have got the data points

schematically indicated like this.   The dose on

this (indicating) particular axis an the proportion

of the study population reacting here.   Obviously,

it goes up to 100 percent.

Then, you have these particular points,
which I have named "ED
50" here. This would be the
dose expected to provoke a reaction in 50 percent
of the study population or 10 percent. This
particular one is an extrapolation, one way of
extrapolating, which one could use.

(Slide.)

DR. CREVEL: What is the impact of the
choice of model on the predicted minimum eliciting
doses? I should go back a bit actually and say
something else.

We collaborated actually with
Dr. Bindslev-Jensen of Denmark in the initial
development of the model. At that particular point
we used the lognormal distribution. Having the
papers published and so on, after that we decided
to go back and look at a few more parameters and
try to refine this particular approach.

Good clinical data were available for egg,
milk and peanut. We fitted the data using the
following statistical distributions and calculated

ED10s, the dose which would be predicted to cause

responses in 10 percent of the population; and

ED1s, in 1 percent of the population for each of

those. We used the following linear extrapolation

from lowest observed adverse effect level to zero

dose, which I showed you that was the red line; the

lognormal model, which was the original one in the

2002 paper; the Weibull model; and the loglogistic

model.

(Slide.)

DR. CREVEL: What I want to do is to

illustrate how these variously fit using the

different distributions. This is using real data

actually from the study by Wensing, et al., in 2002

on roasted peanuts. You have got the data points

here. That is still a normal fit, which is the

original one we used.

You can fit loglogistic pretty well as

well and the Weibull as well. You can even fit a

linear -- you can even correlate these points

linearly as well.

Although I haven't got the parameter fits

here, I can tell you that they are all pretty good

for all of those. Basically, the fit which you use

doesn't actually tell you which is the most

appropriate one for the particular distribution.

(Slide.)

DR. CREVEL: This is illustrated in the

differences between ED10s and ED1s between studies

and models. Now, this is just using data on peanut

actually from -- in this particular case, we are

just comparing the number of studies on peanuts

including studies performed by Bock and May in the

1970s and the later study by Wensing.

What is quite clear actually is for ED10s,

which are still within the experimental zone, what

I call the "experimental zone" in one of the

previous slides.

The data actually drives what the

predictions are. I mean, there is not a lot of

difference between the ED10s, even though it is log

scale, I know, even in this particular case.

(Slide.)

DR. CREVEL: When you move away and go

outside of the experimental zone to the ED1 and

even further actually, the case is even stronger

beyond that, the actual choice of model starts to

drive the predictive responses. That is an

important point to bear in mind.

This is summarized on this slide for the

ED10s in the experimental zone, that the

differences between studies is greater than between

models. In order to address that, the best way of

doing that is to focus on standardizing protocols

and having consistent patient selection criteria

for the studies which you wish to undertake.

For the ED1 values, the differences

between models are much larger as I showed and

increase of course as we move further away from the

experimental zone.

What this actually illustrates is that you

need to validate the particular approach. You need

to validate whatever model you have chosen and

adjust parameters in accordance with that. What

I'm going to talk about is actually how we might go

about doing that, what sort of data is needed.

(Slide.)

DR. CREVEL: Now, there are a certain number of assumptions underlying the values generated by the model. We are talking here about undeclared allergen, and that is quite important in relation to the one of these particular points.

First of all, we assume that the participants in a controlled challenge study are a representative sample of the whole allergic population. That is a very important assumption, and one which actually is sometimes shall we say overlooked.

We have heard a lot about whether people are included or not included in particular studies. I would tend to argue actually that the population used in challenge studies, because of the way they are selected, is actually shall we say at the more severe, more sensitive end of the allergic population, basically because there are people who actually normally are referred to tertiary care centers.

There are people whose allergies are

actually troubling them. They are not people who might just get a small rash and just ignore it or ignore the particular food that caused it. There are people who actually need to manage their allergy and they need some serious advice in doing so.

The second point is actually in terms of validating the model the allergic people actually eat the same foods as the non-allergics. In this particular case, it is quite important because if they are already avoiding them, the number of reactions that you will be able to enumerate in epidemiologic studies will not be the correct one.

The distribution of allergic reactivity study at the population level, now we've heard thresholds for individuals. Minimal eliciting doses for individuals do vary. However, what we are saying here actually is that overall it will be studied in these particular challenge studies.

Finally, the responses to a given dose of allergen are similar in the clinic to those experienced outside. We are doing some work

actually with Jonathan Hourihane, in fact, to try to quantitate the differences that may exist because, in fact, we are very aware that particular assumption probably does not hold entirely.

(Slide.)

DR. CREVEL: Okay. What data do we require for validation and application of a modeling approach? We want to arrive at the risk assessment. We have the hazard characterization. I would put it to you that actually what we are doing by the modeling approach is actually characterizing the hazard. We are establishing how many people are likely to respond to a particular amount, and we use all of these. These particular factors all influence it.

However, we also need to know the number of allergic consumers. That is quite important in terms of prioritizing allergen management and so on. Effectively, what the legislation does is also to acknowledge that particular fact.

The legislation either here, in Europe or anywhere else does not protect everybody because of

course it only specifies a certain number of major common allergies rather than all of the 200 or so foods that may provoke allergic reactions.

We also need to know what the exposure is. We need to know what residual allergen levels are in the foods, residual allergen levels that are not declared. Finally, we also need to know what the number of reactions is overall in the community.

Taking all of that together, we can actually validate the model. Using those sorts of data, we can also apply it properly.

(Slide.)

DR. CREVEL: To summarize, I think the modeling approach complements clinical studies and it certainly compliments clinical studies to establish minimal eliciting doses. Of course, it relies on the data generated in those studies.

I think the advantage compared to just using the data as such is it actually permits more complete use of those data using the whole dose distribution rather than just one particular point, say, the lowest observed adverse effect level or

the no observed adverse effect level.

It also, I think, makes the whole process of risk management more transparent, I guess you would say, allowing a more informed discussion of risk management objectives by all stakeholders. That is very important I think.

In order to agree on objectives, I think people need to know how or need to see the process by which they are reached. However, and this is a big proviso, it does require validation before it can be fully operational.

We are doing work at the moment to see how we can address that. Some of the data actually I should say will contribute to this particular assessment will be generated by some European projects which are currently running, but of course it will take a few years to get there.

That was my last slide. Thank you.

CHAIRMAN DURST: All right. It is open for discussion.

Yes.

DR. BRITTAIN: That was a really

interesting talk. There was one aspect of it that
I'm a little --

DR. CREVEL: I'm sorry? I can't hear you.

DR. BRITTAIN: There is one aspect of it
that I'm a little confused by, and that was in one
of your last slides with all the graphics about the
needing to know the number of allergic consumers.

If you are trying to find the dose at
which the risk of a reaction, given you are
allergic, which is what I thought we were trying to
do, why do you need to know the number of allergic
consumers?

DR. CREVEL: Well, you need to know the
prevalence of the condition within the population.
In fact, perhaps the confusion is there isn't,
because I mentioned validation as well as
prediction in this particular context actually.

For validation, you certainly need to know
how many reactions are occurring in order to see
whether the model actually predicts the numbers of
reaction which you are actually observing.

I mean, this is a big data gap at the

moment. I mean, I don't think either in the U.S. and certainly not in Europe do we have data on actually the number of reactions that do occur. Certainly, we do not have any information on the total number of severe reactions or less severe reactions.

DR. BRITTAIN: You mean the number of reactions that occur across a population as opposed to your study?

DR. CREVEL: Yes. No, across a population, sorry. Sorry, that was in the population, sorry, yes.

CHAIRMAN DURST: Any further discussion or questions?

(No verbal response.)

CHAIRMAN DURST: If not, thank you very much, Dr. Crevel.

Our final speaker for this morning's session is Dr. Steve Taylor. He is the Maxcy distinguished professor and director of the Food Allergy Research and Resource Program at the University of Nebraska, who will discuss Food

Allergen Thresholds.

FOOD ALLERGEN THRESHOLDS

DR. TAYLOR: Well, I would like to thank the Food and Drug Administration for giving me the opportunity to make a presentation to this panel. There are advantages and disadvantages to being the last speaker of the morning. Much of what you are going to see on my slides may just be a reemphasis of some things that have already been said.

I think I got a rather difficult topic, also by being the last one on the agenda, because I'm supposed to talk about uncertainty factors, what are uncertainty factors and how are they derived and what is the underlying scientific rationale for such a factor. I only wish I thought I knew the definitive answers to all of those questions.

(Slide.)

DR. TAYLOR: I think the National Academy of Sciences outlined risk assessment approaches a number of years ago, and I always like to start with this slide, even though I'm not going to

discuss all of these different points, because I

think that the same assessment can be used for food

allergens as is used for pesticide residues and

food additives and other things. This is a very

robust risk assessment approach.

(Slide.)

DR. TAYLOR: I am only going to focus on a

few things on this slide today, and one is

dose/response evaluation. I have been thinking

about this issue for probably 30 years.

This is one of the earliest slides that I

created. At that point in time we didn't know very

much, and I would argue we only know a little bit

more now than we knew when I wrote this slide a

long time ago.

Trace amounts can elicit reactions. I

would argue that the severity of the response is

directly related to the dose. The higher the dose,

the more severe the response.

I would agree that individuals can have

different responses on different days to the same

dose. However, I don't think those responses are
as dramatically different, or at least I would say
that is an unproven point regarding some of the
things that have been said this morning.

There are a lot of assumptions that are
made in this field, and I think as a panel you need
to identify all of the assumptions and question
them.

Stefano Luccioli made a good point, that
individuals vary widely in their degree of
sensitivity in these controlled challenge studies a
millionfold. I completely agree with that. That
is kind of amazing in itself.

The big question is, How much is too much?
The food industry has been focusing on trying to
get an answer to this question for a long time for
some of the reasons that Dr. Crevel just pointed
out.

(Slide.)

DR. TAYLOR: I think there is another part
that we haven't heard quite enough about, and Rene
kind of pointed it out in his presentation. It is

the exposure assessment piece of the equation.

How frequently are food products contaminated with potentially hazardous levels of unlabeled allergens, and how frequently do food-allergic consumers suffer reactions? We really don't know that part very well. Only recently, as Dr. Hefle pointed out, do we have the methodology necessary to determine with any degree of confidence how frequently food products might be contaminated and at what levels.

(Slide.)

DR. TAYLOR: Gil Houben from TNO [The Netherlands Organization] prepared this slide, and I always like to steal good slides from speakers that I invite to be on programs. I think this kind of pictorially describes the situation that exists.

We have food products in the marketplace that contained for one reason or another some level of undeclared allergen. This may be from cross-contact, this may be from use of ingredients derived from commonly allergenic foods that are processing aids and historically haven't been

labeled in most countries.

Then, we have individual thresholds for clinical response that varied by a millionfold as Dr. Luccioli pointed out. There is an intersection here between products that have enough undeclared allergens that at least the most sensitive individuals have some probability of reacting to those.

If I was going to draw this slide myself, I would lengthen the tail of this curve because we know from analytical studies that there are products in the marketplace that are quite hazardous for these individuals containing comparatively higher levels of allergens that provoke severe reactions. Dr. Hefle showed some of those data today.

(Slide.)

DR. TAYLOR: I wanted to say just a little bit about the different kinds of clues that we can have for determining allergen thresholds. Stefano already pointed this out, too.

Probably the best data we have comes from

double-blind, placebo-controlled food challenges or

clinical threshold experiments using double-blind,

placebo-controlled food challenges and

immunotherapy trials that also use challenge data.

(Slide.)

DR. TAYLOR: I actually don't think that

allergen cross-contact episodes turn out to be very

useful in determining thresholds, and I wanted to

emphasize that point, because there is a lot of

anecdotal material in the clinical literature about

these cross-contact episodes.

A lot of them are deficient, because the

analytical methods used to detect the residues in

those studies were probably not as accurate as the

methods that Dr. Hefle described in her

presentation, the methods that we have had for the

last few years. There is often a lot of lacking

information in the investigation of these studies.

(Slide.)

DR. TAYLOR: As I pointed out, this

question of how much is too much has intrigued out

group for a long time in the food allergy research

and resource program.

I want to point out that we are funded by the food industry. We have more than 40-member companies scattered around the world. We began to focus on the threshold question in earnest in the mid-1990s and beyond.

(Slide.)

DR. TAYLOR: We have held a series of threshold conferences. The first one was held in 1999. I was asked to say a little bit about these, and it is really hard to summarize it in 15 minutes or less.

I will point out the fact that the results of the First Threshold Conference have largely been published in the peer reviewed, scientific literature.

The question we asked at the First Threshold Conference is we invited a number of clinicians from around the world to come to South Carolina, because we thought that perhaps they had information on low-dose challenge trials.

When you hear studies of the kind that

Dr. Wood reported this morning, recognize that most

diagnostic challenges start at 400 to 500

milligrams.

No wonder some people have severe

reactions at those dose levels, because those are

quite high in my opinion. We were interested in

clinicians who sometimes, because of the patient's

history, started the challenge at a much lower

level.

(Slide.)

DR. TAYLOR: What did we find out? We

found out that there was considerable data on

low-dose challenges for peanut, egg and milk in

particular and more scattered data for some of the

other foods.

The data were really hard to evaluate

because of the lack of standardized protocols. I

will come back to that in a little bit. The lowest

provoking dose -- we had 306 patients for peanut,

281 patients for egg, and 299 for milk. These

physicians brought this data to this conference.

(Slide.)

DR. TAYLOR: The lowest provoking dose for peanut was about 1 milligram of peanut, which is .25 milligrams of peanut protein. However, I have to tell you that Dr. Hefle and I spent an entire weekend in the conference room trying to figure out what the doses were in these challenge trials, because the physicians don't calculate that, particularly carefully in some cases.

Our personal favorite is the physician that used a drop of peanut butter as his lowest dose. We had him send us his dropper bottle and we tried to figure out how much that actually was. These data look really finite when you show them this way, but there is a lot of glorified guesswork. I just want you to understand that.

(Slide.)

DR. TAYLOR: We determined that minimal eliciting doses or threshold doses do exist for commonly allergenic foods, that the threshold doses are finite, measurable and above zero.

However, it was really difficult to reach consensus, and we didn't reach consensus. We had

about 20 clinicians at this conference, and we did

not reach consensus on what threshold doses should

be.

In fact, for most of them this was their

first introduction to this concept. We had to

teach them what NOAELs and LOAELs were. They make

risk assessments every day but not these kind.

(Slide.)

DR. TAYLOR: We also found that reactions

occur to hidden or undeclared allergens in foods.

No big surprise there. However, severe reactions

to undeclared allergens tended to occur at higher

dose levels.

We also determined that at least in these

populations with these low-challenge doses that

low- or very low-dose exposures, LOAELs, result in

mild reversible symptoms.

(Slide.)

DR. TAYLOR: The Second Threshold

Conference was held in 2002 and was geared to

address the biggest concern we had from the first

one, and that was a lack of a consensus protocol.

(Slide.)

DR. TAYLOR: I don't have time to describe the consensus protocol other than to indicate that it has been published; it does exist; and there are ongoing low-dose challenge trials underway around the world using this protocol or slight variations of it.

As Dr. Luccioli pointed out, most of those haven't been published yet because it takes a year to two years to do these studies to find the number of subjects to enroll in these studies.

(Slide.)

DR. TAYLOR: We did have the Third Threshold Conference where we tried to determine what you do with the data once you collect it.

(Slide.)

DR. TAYLOR: I won't go into that very much, because much of it relates around the modeling stuff that Rene Crevel already described. Because the binomial approaches are just plain difficult, because it is very difficult to identify even 29 soybean-allergic individuals in the world

to do a challenge trial. Believe me, we've been
there, and we know how hard it is.

It is easier to do peanut trials than
perhaps others. It is hard to do milk and egg
because young children outgrow their allergies, so
you've got to be concerned that the child, the
patient, still has the allergy that you are looking
for.

(Slide.)

DR. TAYLOR: There were a number of
advantages to modeling. I think Rene pointed those
out. I will just make the point that the consensus
of the group was that you could do modeling. Of
course, you've got to figure out which model you
are going to use.

Maybe we haven't validated them yet so we
don't exactly know; however, using this lower
confidence interval as the threshold might be a
reasonable approach to consider.

(Slide.)

DR. TAYLOR: Well, classical risk
assessment involves determining the NOAEL for a