UNITED STATES DEPARTMENT OF AGRICULTURE

FOOD AND DRUG ADMINISTRATION

FOOD ADVISORY COMMITTEE

MEETING ON INFANT FORMULAS

Monday, November 18, 2002

8:20 a.m.

U.S. Department of Agriculture
Animal and Plant Health Inspection Service Building
4700 River Road
Riverdale, Maryland

Temporary Voting Members Present

James Anderson, Ph.D.
Robert D. Baker, M.D., Ph.D.
Margaret E. Briley, Ph.D., R.D., L.D.
Scott Denne, M.D.
Cutberto Garza, M.D., Ph.D., Chairman
James E. Heubi, M.D.
Laurie J. Moyer-Mileur, Ph.D., R.D., C.D.
Virginia A. Stallings, M.D.
Patti Thureen, M.D.

Participating Food Advisory Committee Members Present

Annette Dickinson, Ph.D.
Goulda Angella Downer, Ph.D.
Lawrence N. Kuzminski, Ph.D.
Madeleine J. Sigman-Grant, Ph.D.

Acting Industry Representative

Roger A. Clemens, Dr.P.H. CNS FACN

FOOD ADVISORY COMMITTEE ON INFANT FORMULAS

AGENDA

P R O C E E D I N G S

DR. TAYLOR: I'm Christine Taylor, and I'm director of Office of Nutritional Products Labeling and Dietary Supplements at FDA's Center for Food Safety.

We will be this morning going through a series of conversations, and what you've got right now is our overview. So given the fact that we started just a few minutes late, we'll go ahead and try to shortcut this overview.

In the next few minutes, we'll do a brief overview for this Food Advisory Committee meeting. We'll review a couple of the administrative issues, which will be focused on ethics and conflicts of interest. We'll spend some time on regulatory context for this meeting, and then we will begin the meeting per se with Dr. Bert Garza serving as chair.

During the meeting, there will be a series of presentations and white papers, which we have provided as background information for the committee. There will be some discussion, and then beginning tomorrow there will be public comments, more discussion, and response.

We're going to spend just a very few minutes this morning going over the role and expertise of the task force members, and I think the key point to be recognized is that there are several different kinds of members sitting with us today.

FDA considers a number of factors in selecting individuals to serve on the Food Advisory Committee, including their scientific expertise, as well as issues related to conflict of interest.

We also have sitting with us a consumer representative. This person is a voting member of the Committee and represents the consumer perspective on issues and actions that come before the Committee.

We also have an industry representative sitting with the committee. This is a nonvoting member, and they're responsible for representing all members of the industry, and not any particular association, company or product.

Basically, the kinds of members we have on this committee are temporary voting members, as well as some members of our larger Food Advisory Committee. As I mentioned, we also have consumer reps and industry reps.

What I'd like to do, at this point, starting with Dr. Baker, if you will, Dr. Baker, just so we can get used to using the microphones, if you would let us know your name, where you are, and if you're too modest, I have a listing of your expertise.

DR. BAKER: Robert Baker from Buffalo, New York. I'm a pediatric gastroenterologist, and I have a Ph.D. in biochemistry and in nutrition.

DR. STALLINGS: I'm Virginia Stallings, from Children's Hospital in Philadelphia. I'm the head of the Nutrition Section there, and I do work in healthy children and children with chronic disease related to nutrition.

DR. HEUBI: I'm Jim Heubi. I'm a pediatric gastrologist, as well. I'm the program director for the GCRC, the General Clinical Research Center at the Children's Hospital in Cincinnati, and I have a longstanding interest in nutrition relating to infant nutrition bone disease, cholesterol, metabolism, you name it, there's a variety of things.

DR. ANDERSON: I'm Jim Anderson. I'm at the University of Nebraska Medical Center in Omaha, Nebraska. I'm chairman of the Department of Preventive and Societal Medicine, and I'm a biostatistician by training.

DR. DOWNER: I'm Goulda Downer, a doctorate in Human Nutrition, with a residency in pediatrics at Georgetown. Currently, I'm a clinical nutritionist with my own practice, and I'm also on faculty at George Washington University.

MS. SIGMAN-GRANT: I'm Madeleine Sigman-Grant. I'm a maternal and child nutrition specialist at the University of Nevada Cooperative Extension.

DR. MOYER-MILEUR: I'm Laurie Moyer-Mileur, from the University of Utah. I'm a registered dietician with

a doctorate in exercise physiology, and I have over 20 years of neonatal nutrition experience.

DR. GARZA: I'm Bert Garza. I'm a professor of nutrition at Cornell University. I'm both an M.D. and have a Ph.D. in nutritional biochemistry and metabolism, and my primary interests have been in maternal-child health, with interests in growth, and protein and energy metabolism.

DR. KUZMINSKI: I'm Larry Kuzminski. I'm from Duxbury, Massachusetts. I'm retired from the food processing industry, having R&D responsibilities and operations responsibilities with the Kellogg Company and with Ocean Spray Cranberries.

DR. DENNE: I'm Scott Denne. I'm from Indiana University. I'm a pediatric neonatologist. I have a longstanding interest in neonatal nutrition, specifically, and protein and energy metabolism.

DR. THUREEN: I'm Patti Thureen, a neonatologist from the University of Colorado in Denver, and my particular interest is in protein and energy metabolism in the extremely low-birth-weight neonate.

DR. BRILEY: I'm Margaret Briley from the University of Texas at Austin, and my expertise has been in nutrition of children and child care.

DR. TAYLOR: If we could just stop right there. Margaret is our consumer rep, and on our right we have

Dr. Roger Clemens, who is substituting for Annette Dickinson, who is our industry rep.

I'll go to the next slide and just give a minute or two about the staff you have sitting at the table with you. As I've mentioned, I'm with the Office of Nutritional Products Labeling and Dietary Supplements.

We also have Dr. Susan Walker, who is our associate director for Clinical Affairs, as well as Dr. Beth Yetley, who's the lead scientist for nutrition.

Jeanne Latham, who is sitting next to Dr. Bert Garza, is our executive secretary, and we're being joined today by Ms. Mary Ann Killian, who is program integrity adviser at the Office of Human Resources at FDA.

Let me just spend a very quick minute, and then we will return with a regulatory context. I think in terms of mechanics, we need to understand kind of where we are in the process. Currently, we are operating as an ad hoc task force to the Food Advisory Committee. In the very near future, we will constitute an Infant Formula Subcommittee of the Food Advisory, but currently we are still in the ad hoc mode.

The current focus of the Infant Formula Advisory Meetings is to obtain scientific input for evaluating whether new infant formula supports normal physical growth of infants. This comes under Section 412 of the Food, Drug, and Cosmetic Act, which in a few moments

we'll come back to in more detail. We're looking basically for scientific input, which eventually will inform the Agency relative to regulatory efforts.

What we're undergoing currently is a series of meetings, and I'm sure most of you remember that last April we held our first meeting on this issue of normal physical growth. It included a somewhat general discussion, an effort to understand the regulatory context, as well as a few specific questions about extrapolation and attrition in the study.

This is the second of this series, and the general scientific topics for today fall into three categories: Growth Measures and Methodologies, the Role of Such Measures and Methodologies in Demonstrating Normal Physical Growth, and then, finally, Principles and Criteria to Determine the Need for a Clinical Study to Provide the Agency an Assurance of Normal Physical Growth.

It's always helpful to be clear about what's not on the table. There are so many issues in the area of infant formula, normal physical growth, other issues related to the Agency's regulatory purview that sometimes it's important to realize there are things that are of great interest, but are not on the table for discussion.

This lists a few, probably the ones that our discussions will most likely tend to gear toward. The

design and conduct of studies is not on the agenda today, other endpoints of clinical studies is not on the agenda. What constitutes major and minor changes is not on the agenda. That's, of course, for those of you that are intimately involved in the regulatory, you understand that that has regulatory meaning. The nutritional impact or efficacy of formulas, the safety of individual ingredients and specific regulatory decisions are not topics for today.

We have provided specific background for the committee in the form of white papers. We have a total of nine white paper which, as Dr. Garza will explain in a few moments, we'll go through this morning. Each of the papers will be introduced by an expert, and then of course discussed by the committee as appropriate. Those related to the assessment of normal physical growth are listed here, and then for our second topic, changes warranting a clinical study, we have two white papers. Those should be in your notebooks and available for further discussion.

Now, the Agency's role is to give you specific charges that are to be accomplished by the end of the meeting on Tuesday, and those charges are in your notebook in the form of seven questions, and I won't go through them now. I think Dr. Garza will take the time to do that with you later on, but they fall into

basically four categories: Metrics for evaluation of growth, which is Questions 1, 2, 3A and 3B; questions about comparators, Questions 4 and 5; controlled feeding parameters, Question 6; and then changes in composition, Question 7.

Just for the group of us here, the summary of the charges fall into two categories: The criteria for adequate evaluation of normal physical growth during the first six months, and here are several substantive ones. Again, they are specifically it in your questions; and then, secondly, the type of changes in infant formula that should warrant a clinical study. Again, those are the remaining questions in your notebook.

Just in terms of the mechanics, this morning, next, we'll cover the administrative issues, Jeanne Latham, as assisted by Mary Ann Killian, will go through that with you.

I'll return, and with the help of Dr. Walker and Dr. Yetley, give you some regulatory context and then the actual task force meeting will begin.

What we'll do is hold questions until after the administrative component, and then again after the regulatory context, and then we should be on our way.

So, Jeanne, I'll turn the meeting over to you, and dutifully return for the next part. Thank you.

MS. LATHAM: Good morning. I'm Jeanne Latham and, first of all, in terms of administrative issues, we wanted to have Cathy DeRoever's statement read into the record, and Dr. Garza will take care of that.

Thank you.

DR. GARZA: Catherine DeRoever, the executive secretary of the Food Advisory Committee, was asked to take a few minutes to refresh everyone's memory about a few of the rules of the road, in terms of Advisory Committee operations, so I'm going to be reading her statement.

It is my understanding that all committee members have been provided with a copy of a Committee Member Guide to FDA Advisory Committees and a video. The video's title is "A Panel Member's Responsibility." I believe there are copies of the Member Guide available at the registration desk for anyone who may be interested. The Committee Member Guide is in need of updating but, by and large, it provides a good operational overview.

FDA relies on its Advisory Committees to provide the best-possible scientific advice available to assist us in making complex decisions. Our goal is to do this in as open and transparent a manner as possible. Part of that openness carries with it a request that the members try to avoid even the appearance that issues are being

decided or conclusions are being reached outside the actual meeting.

We understand that issues raised during the meeting may well lead to conversations over breaks or during the meal. In fact, we hope the discussions are thought-provoking. We have had instances where the members have come back from a break and said, "You know, we were talking over break, and we would like to request that FDA provide us some additional information so we can better understand thus and such." This is perfectly acceptable.

What we don't want is to have a situation where after the break the members come back and say, "We were talking over break, and we decided that the answer to Question 1 is..." From our perspective, that would be particularly troublesome because neither the Agency, nor the public, would have had the benefit of listening to the entire discussion, the questions raised, the responses, et cetera.

In fact, FDA has recently adopted a policy that only matters that can be decided by a show of hands are procedure matters, for example, break times. I'm not sure I understand that.

[Laughter.]

DR. GARZA: All other votes and comments must be placed on the record, attributed to the member making the

statement. The policy goes even further. If a member has to leave the meeting early, that member waives the right to vote. You may wonder why would the person lose their right to vote, but the answer is fairly simple. FDA believes all parts of the meeting and the discussions are important. Consequently, voting on issues without having the benefit of all of the discussion would be premature.

The issue of openness is larger than what transpires during the course of the meeting. I would like to call your attention to the section in the Members' Guide, titled, "Member Interaction Before, During and After a Meeting." In essence, this section underscores the fact that all communication with the members should be routed through the Committee's executive secretary. No one, not even FDA staff, with the exception of the executive secretary, should be contacting the members about upcoming meetings, topics, et cetera.

This same guidance applies to consultations between members prior to a meeting. If a member receives an inappropriate contact, the members should feel free to notify the executive secretary and/or refer the person making the contact to the executive secretary. Our goal in having all contacts routed through the executive

secretary is to minimize any situation that could be misinterpreted.

Appearance issues are always difficult because, as is true of many things, appearances can be deceiving. We ask that our members, guest speakers, and everyone attending the meeting be mindful of how an interaction between a member and anyone, for that matter, might be perceived.

Please let me be clear it is not my intention to question anyone's motives or integrity, but I am very sensitive to the issue because I have, and imagine so have you, seen highly respected individuals become the object of negative attention based on a misperception, and I certainly wouldn't want anyone in this room to become such a target.

I am confident that everyone here is sensitive to these issues and can appreciate that my comments are intended as a gentle reminder.

Thank you.

Any questions? Which I will refer to Ms. Latham.

[Laughter.]

DR. GARZA: From any of the committee members? Is all of that clear?

[No response.]

DR. GARZA: Thank you.

MS. LATHAM: Good morning. I am Jeanne Latham, the executive secretary for the FDA's Food Advisory Committee on Infant Formula. I want to welcome everyone, and I'd like to read the conflict of interest statement for the record.

The following announcement addresses the issue of conflict of interest with respect to this meeting and is made a part of the record to preclude even the appearance of such at this meeting.

By the authority granted under the Food Advisory Committee Charter of July 2002, the following individuals have been appointed as temporary voting members by Joseph A. Levitt, director, Center for Food Safety and Applied Nutrition:

James Anderson, Ph.D; Margaret Briley, Ph.D.; Robert Baker, M.D., Ph.D.; Scott Denne, M.D.; Cutberto Garza, M.D., Ph.D.; James Heubi, M.D.; Laurie Moyer-Mileur, Ph.D.; Virginia Stallings, M.D.; Patti Thureen M.D.

The issues to be discussed at this meeting are issues of broad applicability. Unlike issues in which a particular sponsor's product is discussed, the matters at issue do not have a unique impact on any particular product or manufacturer, but rather may have widespread implications with respect to all infant formulas and their manufacturers.

To determine if any conflicts of interest exist, the committee participants have been screened for interest in companies that make infant formula. As a result of this review, in accordance with 18 United States Code, Section 208(b)(3), Dr. Cutberto Garza has been granted a particular matter of general applicability waiver that permits him to participate fully in the matters at issue. A copy of the waiver statement may be obtained by submitting a written request to the Agency's Freedom of Information Office, Room 12A30 of the Parklawn Building.

With respect to FDA's invited guest speakers, there are reported interests that we believe should be made public to allow the participants to objectively evaluate their comments.

Dr. W. Cameron Chumlea has a grant from Nestle to serve as a coordinating center for a nutritional study of Chinese elderly.

Dr. Samuel Fomon previously consulted with firms that make infant formula and is likely to do so in the future.

Dr. Duane Benton owns stock in Abbott Laboratories, and he receives retirement benefits from Abbott.

Dr. Dennis Bier's employer, the ARS Children's Nutritional Research Center, recently received the

Bristol-Myers Squibb-Mead Johnson nutritional 2002 unrestricted nutritional research grant. As Center director, Dr. Bier is named as the principal investigator, although no funds come to him personally or for his personal research.

We would also like to note for the record that Dr. Roger Clemens is participating in this meeting as the acting industry representative and a nonvoting member of the Committee.

In the event that the discussions involve any other issues not already on the agenda, for which FDA participants have a financial interest, the participant's involvement and their exclusion will be noted for the record.

With respect to all other participants, we ask, in the interest of fairness, that they address any current or previous financial involvement with any firm that makes infant formula.

Thank you.

With that, I will turn the program back over to Dr. Taylor.

DR. TAYLOR: Thank you very much, Jeanne.

Our goal for the next 15 or 20 minutes is to set the regulatory context for the discussions we're having today. For those of you that remember the spring meeting, we did spend some time on that, and hopefully

most of this is a review, and all we have to add is an additional focus relative to the topic for today. As I've mentioned earlier, this is an ad hoc task force of the Food Advisory Committee, addressing infant formula issues.

Obviously, we have statutory authority relative to infant formula, and its long history goes back to 1980, at which time Congress passed special legislation that amended the Food, Drug, and Cosmetic Act. We, in the Agency, try to avoid throwing numbers, and clauses and phrases around, but it's almost impossible not to, and the key phrase is that it provided Section 412 to the Food, Drug and Cosmetic Act.

In 1986, Congress had an interest in adding to this, providing more statutory authority, and so there were some additional amendments in 1986.

I think what we have to keep in mind is that the infant formula legislation happened for a very specific reason. Infant formula is unique from other foods. It is the sole source of nutrition for a vulnerable population. In Congress's mind, it therefore warranted a special set of provisions for regulation.

It's clear that the intent and outcome of this action was the following statement from Congress. It should not only be safe, which I would point out is handled in separate sets of provisions than what we're

addressing today and contain all of the necessary nutrients, which again is the separate set of provisions which are not on the table for today, but also should provide those nutrients in a bio-available form to ensure that the infant formula were to support optimal infant growth and health. That's what we're about today in some respects.

This chart is a little complicated at first, but I think it sets the context for what we call to be regulatory boxes, and really what's on the table today is this particular component, but all of this is the regulation of infant formula. The safety of the individual ingredients, the classic safety considerations are handled under a separate set of provisions, the so-called Section 409. So ingredients for intended use, that's where most of your classic safety reviews come in.

Section 412, as provided for by Congress in 1980, is really a statutory check on a particular formulated product. In providing those assurances, companies consider the required nutrients that have to be in the formula, the good manufacturing practices or GNPs and quality controls, and then quality factors.

Today, we are focusing on quality factors. These assurances are provided prior to marketing, and once marketing occurs, in the world of infant formula, the claims then come in, as far as efficacy, truthful,

and not misleading, again, a separate set of provisions. I'll come back to this in a moment, but the key component is that we're here taking a look at that.

Now, as mentioned just a second ago, in order to provide these assurances, manufacturers submit a notification to FDA 90 days prior to their intention to market that particular infant formula. Again, as I mentioned, it's specific to a finished product. The Agency reviews it, again, as I just mentioned, for those three components, and it's here, hopefully, highlighted in red that our questions today will focus.

The definition of quality factors is not precise. They certainly do offer the opportunity of expanding, as needed. There is language from a 1980 discussion in the House Committee, and their references to quality factors focus on things such as pertain to the bioavailability of a nutrient and the maintenance of levels or potency. They discuss at great length the growth of infants during the first few months of life, and they discuss the concept of healthy growth, the idea being that once you've formulated a product, it needs to support healthy growth. So, in its simplest form, quality factors are a check on the concern that once you get the entire product put together it works appropriately.

Now the types of quality factors could be many. At this point, we basically have two. In the realm of nutrient-specific, we have provisions for protein efficiency ratios, protein per se, but over time, others could be put in place. In the world of the formulation itself, the totality of the formulation the quality factor we address is normal physical growth, and, again, others could be put in place over time as needed.

So, for today, normal physical growth as quality factor is what's on the table.

The scientific questions that will come through as you read the charges are basically twofold. How do you measure and affirm normal physical growth and how and when should assurances of normal physical growth be appropriately provided?

Going back to that again, this particular slide, quality factors, normal physical growth, assurances for a specific product, along with other components of these assurances.

Now, just so that we're sure how it works from a regulatory perspective, we've put in this slide, but I think it's redundant to what we've said before. In order to provide assurances, vis-a-vis Section 412, the manufacturer submits a notification 90 days prior to marketing. FDA reviews the notification package, taking into all of the components, nutrients, GNPs, quality

control and quality factors, and if assurances are adequately provided, FDA does not object to the marketing of the formula.

If, in the Agency's opinion, assurances are not adequately provided, FDA does let the company know that it objects to the marketing of this particular formulation. It's important to note, from a regulatory perspective, that this is not a premarket approval process, so manufacturers do have the right to go to market over FDA's objections.

The scientific input we get from you folks today and tomorrow will certainly guide our thinking about the evaluation of normal physical growth when infants are fed a new formula. We'd like to point out that it certainly helps us, but it's also helpful to stakeholders in that what is expected becomes clear to them. It's not as much of a black box if it's quite clear how FDA's scientific considerations are handled.

It's also going to guide our thinking about when clinical studies should accompany formulation of processing changes in infant formulas, and again it's helpful to us, but it's also helpful to our stakeholders.

The outcome of today's discussions can be used to inform our ongoing reviews, but we do need to talk a little bit about current rulemaking, in that discussions today have the opportunity or the possibility of

impacting on current rulemaking activities. The current state of our rulemaking, as probably many of you know, is that in 1996, we proposed a rule to implement parts of Section 412, and in that was included the implementation of quality factors.

That final rule has not been issued, so we are still in the process of what's known as rulemaking. If input from this Committee is relevant, and it may or may not be, but if input from this Committee is relevant, there would need to be an opportunity to comment on that, and we would, of course, reopen the comment period on this rule for that purpose. So we retain the option of reopening the comment period.

So, again, just to review, you've seen this before, today's discussions, vis-a-vis the charges, our growth measures and methodologies, the role of such measures relative to normal physical growth, and the general principles and criteria to determine the need for a clinical study to provide assurances of normal physical growth.

I think we've gone over topics not under discussion, so I'll mention these only in passing, and then again remind you that the specific charges from the Agency to the Committee are the seven questions in your notebook, and I'm sure Dr. Garza will go over those with you in some detail.

I do want to introduce Dr. Susan Walker, who's at the table here, our associate director for Clinical Affairs, and Dr. Beth Yetley, who's our lead scientist, and the agreement we have is if you have questions, I will go join them, and we will answer them as a troika.

Thank you very much.

MS. LATHAM: Are there questions?

DR. GARZA: Are there any questions to Dr. Taylor?

[No response.]

DR. GARZA: Thank you very much. That is clear, judging from the lack of questions from the Committee. I do want to take this opportunity to welcome the Committee members, and guests, and staff that have joined us, and to thank Dr. Taylor because we've made up the lost time. I was concerned that we would be running late, but we're doing all right in terms of time.

We have a very full agenda and would like to begin by asking the Committee members if, in fact, they have any questions about the agenda.

[No response.]

DR. GARZA: Very quickly, just to review the procedure, we're going to be launching into one of the major segments of the Committee meeting in just a few minutes, and that is nine presentations, based on the nine background papers which were sent to each of us

several weeks ago. Only Committee members have the privilege of asking questions to any of these presenters. We will try to hold presentations to about 15 minutes. Someone will be helpful in alerting the speakers when I think there is about three minutes left in their presentations to help them allocate their time appropriately, and then we will have about 10 minutes of questions from Committee members to each of their presenters.

You also will have the opportunity to ask questions of those that make comments in the public comment period tomorrow, and we will have blocks of time then to come to some consensus on the seven questions that you have been sent as well.

It's going to be very important that we address each of those questions carefully, and therefore I'd like to make sure that each of you has a chance to review those questions because I will be proposing time limits to assure ourselves sufficient time to deal with each of them in a way that doesn't shortchange any of them, and so we'll be trying to deal with that time allocation later this afternoon.

All experts, I am told, will be able to stay throughout today and tomorrow for those questions and answers to that if, in fact, in those blocks of time when we're dealing with any of those seven questions, any of

the committee members would like to address any questions to any of the presenters, and then obviously that's going to be possible as well.

You have seven questions. They have been divided for us in four sections. One of those sections is on metrics for the evaluation of normal physical growth, a second section deals with comparators for the evaluation of normal physical growth, and a third is on controlled feeding comparators, and the fourth is on changes in infant formula composition.

Rather than reading each of the seven questions, I'd like to take just a few minutes to ask committee members if you have any questions about the issues that we've been asked to consider under each of these sections so that, in fact, we can be clear what we're being asked to do, and you can have those clearly in mind during the presentations.

So let's begin with the first in terms of metrics for the evaluation of normal physical growth. There are two questions under that section. Do any of the committee members have any questions about points that you're being asked to address? I'll give you a few minutes to review those, and we can ask the troika to clarify those for us. I think that's the way you were described. That was not my word.

Having had the pleasure to work with them, you'll get informative responses to your questions, I'm sure.

MS. LATHAM: At the end of last week--this is Jeanne Latham, the exec sec--we e-mailed to everyone the updated questions, and I just wanted to make sure that you all have those, and if you don't, we will get them to you. You've got them. . Anybody that doesn't have them?

[No response.]

DR. GARZA: I think they were in the packets today, again, in case you didn't bring them with you.

Would any member of the staff want to address any questions in this section?

[No response.]

DR. GARZA: I take it, then, that they're clear--Dr. Thureen?

DR. THUREEN: Yes, I have one question. These are metrics for the evaluation of growth between birth and six months of age. We will be dealing with both term and preterm infants, I presume. Should they be handled separately? Because the preterm infants we are discussing I believe growth after post-conceptional age birth to six months, so should they be handled separately or should we just do a general assessment of these evaluations, presumably for term infants, with maybe later adjustments for preterm infants?

DR. GARZA: No, I would assume that we will take those separately, but let me ask the staff if they would object, if there's any reason why we shouldn't take them separately?

Committee members? So we'll probably do A and B. Thank you. That's a good clarification, with preterm being all preterms, low birth weight, very low birth weight, and extreme low birth weight.

Any other questions, then, on this first section?

[No response.]

DR. GARZA: Then, on the second, on comparators for the evaluation of normal physical growth, let me give you a few minutes to review those two questions and see if there are any issues that need to be clarified.

[Pause.]

DR. GARZA: Any questions on either of those?

We have one question on the control feeding comparators. Let's take a look at that and see if that's clear.

DR. ANDERSON: This is Jim Anderson.

I wonder if I could get a clarification of the difference between the current infant formula plus new ingredient that's listed on the first bullet and the infant formula plus new ingredient with the asterisks on the last of the bullets.

DR. GARZA: On the last bullet, right? So it's listed below are examples of controlled feeding clinical comparators, and I believe the question is can you clarify the distinction between the first and the last bullets.

DR. TAYLOR: We're working on it.

DR. GARZA: I gather what it meant was that it was a study in which the new ingredient would be used with some infant formula, but that the intention was to market the new ingredient as a component of some other infant formula.

DR. GARZA: That was mine or a generic infant formula, where that new ingredient might be added to any formula, so that it would be a generic comparison was the way I read that. Am I not clear of that? If my interpretation is correct, so it's a generic--

DR. WALKER: Your interpretation is correct.

DR. GARZA: So it's a generic endorsement of the ingredient.

DR. WALKER: Right.

DR. GARZA: Does that clarify it?

Dr. Walker, would you--

DR. WALKER: The instance in which the generic ingredient is added to the new infant formula and the way that you discussed there, Dr. Garza, is a correct interpretation.

DR. GARZA: So it's more an endorsement of the ingredient itself, with a test formulation, but as opposed to a specific formula that had been marketed in the past, where a new ingredient would be added. I understood it to mean a more generalized evaluation, rather than a specific one.

DR. WALKER: I think after we have some of the discussions, some specifics of these will be made much more clear. I think the speakers will address a lot of these issues in detail, and then we can have more questions.

DR. GARZA: And I've been reminded that we must each identify ourselves before we ask questions. I will try to remember, but I'm probably the guiltiest of all. I would hope that they would recognize my voice before this meeting is over.

The statement I was supposed to read said, "I am Cathy DeRoever." I thought I better not say that.

[Laughter.]

DR. GARZA: It would confuse people, and obviously embarrass Dr. DeRoever as well.

Are there any other questions on this third section?

[No response.]

DR. GARZA: If not, then, the fourth one is the last. It is also one question with two parts, A and B.

The table that is attached, obviously, is quite informative, so I would ask you to take a few minutes to look at Question 7, along with Table 1, see if there are any questions.

[Pause.]

DR. GARZA: Are there any questions related to this section?

[No response.]

DR. GARZA: If not, what I propose is the following; that you think about 30 minutes for each of the first six questions, approximately 120 minutes for the seventh question. If we don't need the entire two hours for the last question, we can always come back and address issues that perhaps we might have felt were not completely resolved.

Obviously, if it's clear that we need more time with any specific question, as the discussions evolve, then we can always go back and try to reallocate them, but we will come to some agreement on how much time to spend with each before we start the discussion this afternoon. But as you've had a chance to review them, see if, in fact, you feel comfortable beginning with that type of allocation for the various questions that we're being asked to review.

I don't think that all first six would necessarily take 30. Some may take a bit longer, and

some will take less, but I want to make sure that we don't shortchange, as I said, any specific question, and so we can get agreement on that as a group, going through the discussion and developing some information that would be useful to the FDA I think is more likely.

Are there any other questions regarding procedures or the charge to the group that either staff or I have failed to clarify for you?

[Pause.]

DR. GARZA: If not, then why don't we begin with the presentations. I know that Dr. Chumlea was on the bus, so I assume he's here. We're starting a bit early, but I think that's fine. I will just introduce each speaker as they come forward.

Dr. Cameron Chumlea is a Fels Professor at Wright State University School of Medicine in Ohio. And for those of you in the field of anthropometry, I don't think Cameron needs any introduction. For those that may not be familiar with his work, we don't have time.

[Laughter.]

DR. GARZA: It has been quite extensive, and he's certainly recognized throughout the world for his work in this area.

Thank you very much for the white paper and for joining us this morning, Cameron.

DR. CHUMLEA: Thank you very much for the introduction. It's a pleasure to be here this morning, and I hope I can provide the committee with some information that's appropriate.

First, I'd like to just simply recognize my co-author, Dr. Shumei Sun, who I know is familiar to many of you and point out that she has just recently become Wright State's Brage Golding Distinguished Professor of Research.

So, first of all, I'd just like to point out that growth is relative, as you can see from the slide here, it says, "I keep track of my son's growth, which is going up the vertical scale, and my husband's growth. Frank is age 30, 33, 35, 40." So growth goes in various directions for all of us here.

The second thing here is that we're dealing with infants. Of course, there's our perspective of what infants are, but there's also a public perspective, and I saw this checking out the groceries and decided it really deserved a slide because clearly this, to some degree, is maybe the public's viewpoint, which you can clearly see. It's amazing what you can do with PowerPoint these days and some slides that are available.

So what I'm going to do this morning is just basically cover some brief information that's probably familiar to everybody, so that we can just kind of all

come up to speed. Clearly, infancy is a period of rapid growth, and to some extent, this is probably the most difficult group of individuals to measure. The only other group that's equally difficult is at the opposite end of the age range, but this is generally a very difficult group, but it's also easy in the sense that there's also very few actual measurements that can be collected from them that are really going to be useful. Weight, recumbent weight and head circumference are the three that are the most important and the ones that should be taken.

Just to review, weight clearly measures the growth of all body tissues, recumbent length describes the amount of linear growth because we're dealing with both increase in mass and increase in size, and then head circumference reflects brain growth because this is the period of time, the first few years of life, when the brain actually does the majority of its growth.

This is a period of time when body dimensions increase at a greater rate than in any other period in life. Weight increases between birth and six months about 115 percent, length increases about 34 percent, head circumference increases 22 percent on average. The rate of growth in weight ranges from about 1.1/1.2 kilos for boys or girls at one month of age, but then we have to remember that we're on a growth curve here, and then

it, of course, starts to slow down, and by six months it's running around a half a kilo a month for boys and girls at six months of age.

The rate of growth in length is about 3.5 to 4 centimeters per month for boys and girls, and it slows to about 1.5 to 2 centimeters a month for girls and boys at six months. Just as a reference, the adolescent growth spurt between, say, you know, 12 to 16 years of age, the maximum amount there is only somewhere in the neighborhood of 5 to 8 centimeters a year. So here we're looking at 3 to 4 centimeters a month. So they can just put that in comparison because everybody kind of focuses and remembers how much growth their kids did when they were adolescents. They frequently forget that that rate was a fraction of what they were really doing in the first few years of life.

The assessment of status, once we've collected measurements, we really need the measurements to be accurate and reliable, and this is really a very critical point, particularly in this particular age range.

The measurements are really not difficult to take, and there are a variety of mediums in which they're now described. NCHS produced a video at the end of NHANES III that describes these and all of the measurements that were used in NHANES III on that videotape. WHO has an in-house video that describes the

measurements that are being used in their multi-center growth reference study. I don't think that one has really actually been distributed yet, but it is available if you can talk to the right people.

These are also all very similar techniques. They are also probably being what's currently being done in the current NHANES, and they're all the ones that have come out of the Anthropometric Standardization Reference Manual from 1987, by Lohman.

When we collect the measurements, this is a point where frequently things get skipped over, and the reason things get skipped over is because there's lack of time, money, personnel, et cetera, but it's the part that's really very important, in terms of collecting the information because, one, we're going to either plot those on a growth chart or refer to status, and the other aspect is what's very important here is we're going to calculate increments of rate of growth, and there you're compounding your measurement errors.

It really takes two people to measure an infant appropriately. I'm sorry, folks, but that's really the correct way to do it. It can be done with one, but that's going to add to the errors that are going to be involved. You need an examiner who's going to position and take the measurements, and you need a recorder who's going to be writing down the measurements because one

person is holding the infant and trying to do the measurements, and they really don't have time to write anything down, and then they need to switch roles because we need to take double measurements here so that we can get as much information as possible.

Also, what we'd like to do is have the technicians compare their values. One, this just simply catches transposition errors that occur frequently because people write down numbers in different ways sometimes. All of the studies that I've described to you have allowable differences between what the measurement values can be between the technicians, again, to control for errors and just to control for variabilities that can occur.

Did I skip one? How do I go back on this?

[Pause.]

DR. CHUMLEA: Weight. An infant can be weighed alone or they can be weighed while the mother is holding them. It depends a little bit on the situation, but, frankly, I would prefer that the mother hold the infant. You can weigh the mother, weigh the mother holding the infant, take the subtraction and you're going to get the weight. The reason I like the idea is it keeps the baby calm, and it provides for a very stable piece of information.

The infant can be weighed alone, and there are a variety of electronic scales, but pretty much once you take the baby out of the mother's arms and you place it on something else, it starts moving around. Fortunately, the manufacturers of several of the scales now can compensate for this weight so the stability of the measurement is much better than it used to be.

It's best that they be weighed nude. Blankets, et cetera, are available. However, if they are going to be in undergarments, I think NCHS subtracts about a tenth of a kilo from that for the readings, and spring-type scales and beam balance scales are simply not appropriate for use any more. The electronic scales are much more available.

There's a company called Seca that makes some very good scales. We've used those. WHO has a really nice platform scale that actually pares the mother's weight. I don't know exactly what the manufacturer is, but it's a really excellent device. I think it's been specially made for them, but there are a variety of scales that are available for use, but my preference is that the infant should be held with the mother or the caregiver, whoever is there, and then subtract the weight, if that's possible.

This is the one nobody really likes. It's recumbent length. It takes two people to do it, again.

A variety of different pieces of equipment that are useful for doing this.

It requires one person holding the infant's head. It requires the mother, the caregiver standing there beside the infant reassuring them that nothing is going to happen, and it takes a third person then to position to footboard up against the soles of the foot of the infant. You're holding the head so the child is looking straight up in a vertical Frankfort plane, and it takes then another person to hold the legs, both legs, for the infant, if it's very small, and get the length, and by the time they're six months, generally, the best you can sometimes do is grab one leg and try to hold the other one with your little finger and get it. It's not easy to do, and it's one that can be particularly prone to error, and it's also important that the kid stays straight down the table. I think I've covered what's there.

This simply gives you a description. I think this child is about two, the age there, but again positioning the head up against the headboard, keeping the legs straight and keeping the feet straight up in terms of taking the measurement, but clearly it takes two people to do.

Head circumference, it should be measured with an inelastic tape of fiberglass, metal, something like

that that's good. It's really best that the infant is seated in the mother's lap. I don't like it being done with the infant. I've seen that done. I like the kid up in the mother's lap, which requires then the person getting, who's taking the measurement, to get down beside the mother. This allows the mother to cuddle the infant, keep him quiet, and you can slip the tape over their head and get the measurement before they pretty much know what happened to them.

It's placed right across the front of the skull, and it can be quickly moved up and down the back of the skull. The insertion tape is a nice piece of equipment that's useful for doing this until you find the greatest circumference, and then you pull the tape tight, and this is something we have to coach people in because it doesn't hurt the infant. There's no pain involved with it. They feel a little pressure, and they want to kind of shake it off, but it does need to be tight.

And you can see here it's just anchored really right over the kid's eyebrows, worked up and down. Most kids at this age aren't going to have quite this much hair, and so it's generally pretty easy to get this one from them, particularly, again, with the child being comforted by the caregiver.

Now, what are some other measurements that could be taken? Well, there's really a bunch of them, but I

don't think those are really going to be really appropriate in the instance here. Crown-rump length is sitting height. Crown-rump length, I think that was pretty only used in children with special cases. Chest circumference, limb lengths, one that's potentially possible is skinfold thicknesses.

The problem with all of these measurements is they really kind of have a restricted utility in terms of describing normal or healthy growth. They're frequently prone to high measurement errors, and there's really a limited amount of reference data available for all of these measurements pretty much within the age range that we're looking at, six months.

If you're going to do something like a skinfold. The skinfolds you're going to take are going to probably be triceps and subscapular. Now, again, you're bringing two people into the program, if not three, and you've got to go to landmarks, and to do a triceps, you've got to find the midpoint of the arm, which means you've got to measure it, and make that determination, and you can see clearly it's requiring one person is holding the child and the other person is taking the measurement, and then you're going to have to go and take the measurement.

The question comes up there in terms of equipment. Skinfolds are dependent upon the type of caliper you're using. There's two major brands, really.

There's a Lang and a Holtain, which you see here. NCHS and WHO are both using the Holtain caliper. I think the Lang is still used out there, to some degree, but I'm not that familiar with it any more. There are differences there.

Really, if you're wanting to do skinfolds, the question is what do you want to get out of that? And probably what you wanted to get is total body fat, and there's probably, I think Dr. Ellis is probably going to talk about better ways of doing that today than taking skinfold measurements. I don't really feel comfortable in doing it on anybody until they're about two years of age. It's just difficult to do.

Now, there's indices that can be used from the information that's collected. BMI is the one that we all get informed, and just for your information, I'm 28, so you can kind of put that in reference. I always think everybody should, when we have meetings like this, they should always walk around with their BMI on it, so we'll all be honest about this folks.

[Laughter.]

DR. CHUMLEA: The problem with BMI is that, in infants, you've got 25-percent body length is composed of the head, so that throws off the proportionality aspect. The relation of BMI with direct measures of body composition in infants hasn't really been established.

Weight for length is probably a better descriptive indices of relative leanness adiposity within children.

Measurement error is very important and needs to be paid attention to. The catch here is that the error which may be small is actually going to be very large because of the small size of the child that you're measuring. So you really need to pay a tremendous amount of attention to error in measuring infants. Of course, they can have a tremendous impact on the interpretation, if you're going to go growth increments.

We need to get good-quality equipment. Measurements should be taken. If they're taken on a daily basis, the equipment needs to be calibrated. That includes scales. People forget that scales can go out of calibration, and then particularly the technicians need to be trained in a standardized way of taking the measurements.

We need to collect inter- and intra-observer reliability. Quality control is really important, particularly if there's going to be more than one center used to collect information because we need to control for inter-site differences.

Measurement schemes. You need baseline, an interim and a final. I really like something that's going to be getting a measurement at 1, 2, 4 and about 6 months of age is my preference for collecting things,

generally, starting after about 10 to 14 days. With measurements at 1, 2, 4 and 6 months, you're going to get a good accountability of weight measurements over that period of time. Clearly, if you can collect more measurements, the more measurements the better. I'd be very happy to have those.

I'm going to cover just very briefly growth increments, which are going to be calculated from the repeated measures of growth, and there are charts that these can be plotted on from birth to 12 and 3 to 6 months of age, which are examples here.

These are from Fels data, and just contrary to popular opinion, the majority of Fels infants were breast fed for at least three months, exclusively, so that has been reported. We tend to not get a good press on that. I just want to kind of correct that.

In terms of growth velocity data, there's the Fels data. Also, I'd like to point out that WHO is collecting longitudinal data from its multi-center growth reference study, but this data and report from that study has not been available, and that'll probably be given later.

So recommendations, from what I've just described to you, weights should be measured I think at 1, 2, 4 and 6 months. I'd like to see recombinant length and head circumference at the beginning and end because

it just gives you additional information on the quality of the size of the infant. Close attention needs to be given to methodology and errors. Two technicians are really important and reliability data needs to be collected, and use of existing increment charts until the WHO charts are available.

So thank you very much. I'd just like to say, personally, the last time I had to give a paper in front of Dr. Briley, she gave me a B--

[Laughter.]

DR. CHUMLEA: --which is about 25 years ago. So I hope I did at least that good this time.

Thank you.

DR. GARZA: We'll take that up at the break, I guess.

DR. CHUMLEA: Okay.

[Laughter.]

DR. GARZA: Thank you very much, Dr. Chumlea.

Are there any questions?

DR. SIGMAN-GRANT: Sigman-Grant. I have a question. You recommend weight starting at one month, and you talked about the regain from the loss from birth weight. So much is happening in that first month. Why don't you measure it before--

DR. CHUMLEA: I said that, really, as early as, say, 10 to 14 days. I would like to see it done that

way, but within no later than one month of age. So let me kind of restructure that between--

DR. SIGMAN-GRANT: Why not between--why not the first week, instead of 14 days?

DR. CHUMLEA: Well, there's a shift in weight, as far as I'm familiar with, that's supposed to occur after birth, and so I think, just my understanding is, that there's a period of time within the first week or State or local that the infant basically kind of stabilizes after the birth experience. Now, Dr. Fomon, I'm sure, could give you more information on that, if I'm incorrect on that.

The more measurements you can get out of this, the better. I was being, trying to give you what I think is the very minimum that you have to collect there. You can measure them every week. That would be fine with me.

DR. DOWNER: Goulda Downer. I understand, when you talked about examiner variability and possibly downright error, but can you talk a little bit more about why you don't think that subscapular skinfolds and triceps skinfold are important--

DR. CHUMLEA: Useful information?

DR. DOWNER: Yes, at this juncture, because I think it is.

DR. CHUMLEA: The reason I don't like them is this. First of all, they're extremely difficult to

collect in children at this age, so the amount of error that's in the measurement is extremely high. The question I would have is that what information are you going to get out of this particular measurement that you're not going to be getting by bodyweight alone? Because if weight is going up, the skinfolds are going to go up; if the weight is going down, the skinfolds are going to go down.

So the question that you're really interested in is total body fat, and, yes, 90-some-odd percent of total body fat in a child is principally subcutaneous; that if you want to go total body fat, there are now better ways of doing that, such as DXA, that I think are going to give you the information that you really want.

If you do go and collect the skinfolds, then you're faced with some reference values that are useful. There's only two that are out there that are fairly good--what's available from NCHS and NHANES III, and then Dr. Fomon's data on skinfolds. But outside of that, there's really little other reference data that's available.

So I guess if I'm going to, what I want to know is total body fat, and if I want to measure total fat, I'll go measure total body fat with something that's going to give me I think better, and more accurate, and reliable information about the child than I'd get from the skinfold.

DR. HEUBI: Jim Heubi. I don't want to misunderstand what you're saying, but you're not recommending that people weigh infants in garments and subtracting--

DR. CHUMLEA: I'm sorry, what?

DR. HEUBI: You're not recommending that people weigh infants in garments and then subtracting one-tenth of a kilogram for the--

DR. CHUMLEA: I would prefer they be weighed nude, yes.

DR. STALLINGS: Stallings. To follow up a bit on the other question. We're beginning to think I think about looking at infants who are growing too slowly as the historical way of the failure to thrive related to this, but the issues of growing too fast are also of concern.

So, to go back to the question, could we get most of that information with weights and heights, and weights for heights, rather than looking for data related to adiposity, or if we were looking for excess growth, whatever that concept means, what would you recommend, derived values from the anthropometry or DXA?

DR. CHUMLEA: I think--we're still sticking between this birth and 6-month range, and I think if you're getting excess growth, you're going to, you may need to take more frequent measurements so that you can

plot and get a better description of the curve as what's going on there. And if you have more information, then you can discriminate between the children who have excess growth and those who don't. So that's I think something that's important to consider.

I think you will get everything you want from weight and length. If you go to DXA with a child at this age, and Ken will address this more I think in his talk, you're going to get, you know, fat, lean and bone, and the fat is probably the most important aspect here that you'd be concerned about for excess growth, but it's going to be described in weight, also.

So unless you're wanting to tease a tissue out and say, okay, we're really concerned about the increase in fat here, in addition to the increase in weight, then, yeah, then something like DXA I think would be important.

DR. STALLINGS: Follow-up. The velocity, then, would be what we would be looking at, more than just attained weight?

DR. CHUMLEA: I think you have to do both of them.

DR. STALLINGS: We, historically, are always looking I think at the attained weight.

DR. CHUMLEA: Yes, and you'd have to include the velocity in there because the velocity would, these

children should, I think, potentially have much higher velocities.

DR. STALLINGS: And that might be a way of discriminating between the concept of normal growth and excess growth?

DR. CHUMLEA: Yes, right. You could have children, let's put it this way, who have, say, after, say, three or four months, when their velocity should be declining, these children might not be declining as what the average is, so they're still obtaining a rather high velocity of growth at that point.

DR. THUREEN: Thureen. I would argue that DXA is not a very useful body composition measurement for most studies because it's not that readily available, especially as a field tool, and I think that a lot of people are now starting to do more caliper measurements for assessment of body fat, even in very tiny infants. And certain people, like Suda Kashyap, have gotten very reliable measurements over time.

Do you think that the data from the NHANES study on body composition, using anthropometric measurements of body fat, was not useful or do you think it is useful? And if you want to do large population studies, do you think that there is a future for caliper measurements?

DR. CHUMLEA: The DXA thing you can talk to Ken about. I'll let him address that when he gets up here.

There's two issues here, whether we're talking about small studies or large studies. If you're wanting to do large-scale studies, population studies, like NHANES has done with NCHS, then collecting caliper information is going to be what you can do because there can be limits to what you can collect for DXA, particularly because of the issue of radiation exposure, although it's very minimal, even within the current NHANES, where they have DXA machines in all of the trailers, I think the limited age there is Age 8, from what they're collecting, although they technically have the availability of doing it in those particular studies. Other people clearly don't have access to such expensive pieces of equipment.

If you don't have access to that, then I'm not opposed to collecting the skinfold data. The issue comes up that it is extremely difficult to collect accurately and reliably, and so, I guess, I kind of am in favor sometimes of no data is better than bad data, and I know that, in collecting it, it is something that people have to pay very close attention to, the technicians have to be very careful, and this is frequently something that in the course of studies, we pay lip service to it, and there's good attention, but these things do tend to fall out.

Now, in smaller scale studies, this is something that can be done, and the information can be collected. Overall, I'm just not that happy with the information. Yes, there are studies where it's been done very well, and so I'm not putting those studies down at all. I'm just talking about, in general, my experience has been, in collecting from infants in this age range, that this is really hard to do, and when it's something that's hard to do, it doesn't sometimes always get done the best way.

DR. THUREEN: Thureen, one more question.

In your opinion, if you're looking at a growth outcome study, do you think incremental data are the gold standard--growth data are the gold standard, attained growth are both critical to an outcome study?

DR. CHUMLEA: I'm assuming, when you say a "growth study," you're going to be collecting repeated measurements from the same children, so you're going to have both pieces of information available there.

The status value simply describes where the children are in reference to whatever reference values you're using for peers at that age. That simply tells you that they're at certain percentile levels, but at the same time children also grow at different rates, and so there's a distribution of the rates at which they grow.

So children who may appear to be at one percentile level, their rates of growth can be a

different percentile level, so it gives a much clearer picture upon what's available. And since any study where you're going to collect repeated measurements, you're going to have all of that information available to you. So I would take advantage of it. Again, the errors are difficult to control and need to be paid attention to for collecting it.

DR. GARZA: Cameron, I have two questions. Given the fact that we are going to be providing advice to the FDA on the approval of specific formulas, how many measures do you recommend be taken if, in fact, one has an interest in the pattern of growth?

DR. CHUMLEA: If I was going to design the study, and you're not going to restrict me to what I want, okay.

DR. GARZA: From your perspective, if you're going to be protecting the public health and infants' health, what should the American public ask?

DR. CHUMLEA: I'd want a birth weight.

DR. GARZA: Birth weight.

DR. CHUMLEA: Clearly. I would like it at, say, two weeks, one month--I'd like it again at two, and then at four, five, and six maybe, something like that. The more measurements I could get out of the thing the better.

DR. GARZA: But you think that with seven measurements, one would be able to assess both the pattern of growth, as well as velocity of growth, at those specific time periods.

DR. CHUMLEA: Yes.

DR. GARZA: And you mentioned there were various sources of error. Is there any consensus that FDA could rely on that deals with the nature of the equipment, the type of calibration that should be insisted upon, the training that obtaining the measures should be able to obtain and document, and the--those would be the three: equipment, calibration and the training of the technicians.

DR. CHUMLEA: There's a little bit of information about equipment errors and a little bit of information about the inter- or intra-observer errors for collection of measurements in the Anthropometric Standardization Reference Manual that's collated in one location.

There's other pieces of information that are clearly scattered around the literature that are available. From NCHS, there's really a limited amount of information. There was really limited error data that was collected in NHANES III. What's available from NCHS is principally from the earlier NHANES studies, and NHES. So there's not much there.

The techniques are described in a variety of locations, but there's not really anything that I know of that's really written down that says, okay, you can refer to here, and this is what you should do, in terms of training, collecting the measurements, et cetera, in one central location, no.

DR. GARZA: Any other questions or comments?

[No response.]

DR. GARZA: Thank you very much.

DR. CHUMLEA: Thank you very much.

DR. GARZA: We'll move on, then, to the next presentation. It's a topic that has already come up, body composition assessment in early infancy. Dr. Ken Ellis, from the USDA/ARS Children's Nutrition Research Center, with Baylor College of Medicine.

Again, Dr. Ellis, thank you very much for joining us.

DR. ELLIS: Thank you. What I'm going to present today is a probably a little different from what most people have had experience seeing. Some of this--in fact, all of this is going to be body composition beyond simple weight. If one was interested in what the composition of weight is--let's see. Which button do I...this is the laser, okay? And which one of these is--can you hear me now?

[Pause.]

DR. ELLIS: This is supposed to move the slides, she said. I'm so used to pointing to the screen these days, so it's my fault.

As you already heard, most people--or most of the work, at least in infants, has been with weights and heights, all kind of weights and heights charts. But body composition, at least the first models that were attempted, the basic classic 2-compartment model is to measure--is divided in two compartments, fat and non-fat. The direct measurement of body fat is really very difficult to do. It's not an easy process to do that. And so for many years what we did was we said that if we could measure precisely some parameter of the body that would represent the fat-free mass, then subtraction of the fat-free mass from the total weight would give us a measure of the fat mass.

Part of the problem with that is that all the years in the (?) scale for the fat-free mass translate directly to the fat mass, and you'll see what I mean by that in a few minutes.

Three classic methods that have been used for 50 years, or maybe even longer: underwater weighing in adults--you can't do this in babies and infants. Practitioners as well as parents tend to object to holding babies underwater. Hydration, this is probably the more common thing you'll find in the literature that

is done. It's the dilution technique. You give a tracer, collect the blood sample or some fluid sample several hours after that, then do some manipulations on that, (?) space to get water, and make some assumption about how much water there is in the fat-free mass. And as we all know, hydration content of the fat-free mass in children at very early ages changes dramatically. So depending upon what you assume, you can then quickly be off in your estimate of the fat.

Whole-body counting is another method used to measure potassium content, primarily the body cell mass. And, again, how much that relates to the--how much of that is a constant fraction or not of fat-free mass at these ages has also been questioned.

So, again, like I said, the difficult underwater weighings, difficult to do the infants. For the water measurements, they must swallow all the tracer, collect some kind of fluid sample. Plasma is your best choice.

The problem with this is you can't repeat it. If you want to do the trial a week or so later, you can't. You need to leave a sufficient amount of time for the previous tracer to clear, or you start increasing the doses of the tracer to compensate for that. And, again, in the past, the most accurate assays required one to have a mass spec or availability of a mass spec, which is not in everyone's laboratory or garage, as I usually say.

Whole-body counters, the problem with those have been over the years really they haven't been designed for infants. They're really designed for adults. There are a few of us who have done this, but in general they simply don't exist. And even if you do have a whole-body counter available to you, most of the time it's not in a clinical setting. You're going somewhere else to get the measurement done. Again, the reason that tends to exclude it, at least for infants and children. But one nice feature is you can repeat this as often as you wish, and so you could do it on a daily basis if one chose to do that. So if you have access to a counter, one can count these infants as frequently as one chooses.

And because of the way in which the counting procedure works, there doesn't have to be any really significant constraints. They can move around and be--it will not really affect the results that much.

Now, I looked at the various things that we received in reference to this meeting, and there's the document in there from the American Academy of Pediatrics where there's one little paragraph on body composition. This was in June of '88. And it says, "Normal growth implies appropriate composition of the increment in body weight. Sequential measurement of various aspects of body composition"--such as water, fat and bone--"have the potential for defining changes in body composition."

However, at that time the opinion of the Task Force was that such measurements have not yet reached the stage of precision, non-invasiveness enough, and they're not that very convenient, and which I've just showed you that sort of in the three previous procedures, which I would agree in 1988, I would agree with that statement.

That gives us a quick summary of that. It just says you want to be able to--that the weight is appropriate composition, you will be able to do longitudinal measurements. Again, for clinical testing, the precision, noninvasiveness, and convenience are the issues that they put out.

Again, the 2-compartment model, I quickly talked about the limitations. The density is not constant. The hydration is not a constant. Extracellular and water ratio is not constant. Bone accretion is not constant. Basically babies aren't constant. We know that, right? Because we know that when the baby comes back a month later, it's not the same baby we saw a month before. It's a whole different child totally, at least from a body composition point of view, unlike adults, which really change very slowly over time.

Now, so what's happened in pediatric body composition research since 1988? There are sort of three general areas where there have been advances made. One is bioelectrical techniques. In pediatrics, I want to

take a broad sense of pediatrics. That's anybody under the age of 18. And there's been quite a bit of work used with this technique in older children. When it has been tried or used in infants, it has not been very successful at all. In fact, in general, most conclusions with this technique--I should say the first two techniques, the bioelectrical impedance and the bioelectrical spectroscopy, which is the same as this but at two frequencies, have been that the information gained from those two techniques really hasn't been much more than you already knew when you had just simply weights and heights with these individuals.

TOBEC, on the other hand, has been more successful, but the problem with TOBEC is that these machines are very--they're not common, there are not many out there. There's probably not more than a dozen--a half a dozen, in fact, for infants. And so, again, this is a technique that holds promise or has held promise, but, again, it's not a technique that is widely available at all.

Absorptiometric techniques, DXA. Remember that Academy of Science report? It was in 1988. Well, in 1993, the world of X-ray absorptiometry changed because at that time it went to what's called DXA. It went to X-ray sources, different detectors. The whole technology advanced substantially such that one could now consider

this technology for whole-body measurements in infants, and one can do a localized region, such as the spine, if that's specifically what one would want to do.

Another area, body volume measurements. Remember, underwater weighing doesn't work. This is an air displacement plethysmograph technique now, which is just actually started this year, so maybe in another few years, when this committee reconvenes, we'll be able to talk more about this.

I'm going to basically focus on the absorptiometry or DXA or DEXA methodology because that right now holds the best promise for this kind of broad application.

So the basic model, again, in 1988--I mean, there are more models, but the very basic model, again, we're talking about fat and fat-free mass. And today, when one talks about body composition, one is almost forced pretty much to really address this kind of a model over here, 4-compartment model: fat, bone mineral, mineral composition, ash, the water, and protein content. And, in fact, it is these compartments that we're interested in when we look at body composition at any age and look at change in body composition at any age.

So that 4-compartment model is now shown on your left, and what does DXA provide us? DXA provides us--the only method we have that provides us a 3-compartment

model for a single measurement. We have a measurement of fat, which is directly the fat. We get a measurement of BMC, which is bone mineral content, which is for the mineral compartment. Eighty percent of this is this; there's another 20 percent which is distributed in the non-mineral, non-osseous compartments. Then a third large compartment called the lean tissue mass, which is the non-bone, non-fat compartments.

So when we do a DXA measurement of anyone at any age, we get this basic model, which has been a huge advance.

This is what an image looks like in a child. Most of the time people show images of adults. Adults are very nice. They hold their arms right. They put their legs down straight. Children, amazingly, 14 years of age, haven't learned how to do that yet, have they? So they tend to lay the way they want to. We do make measurements in children at our place. We've done probably, I don't know, 600, 700 DXAs, at least. Maybe a thousand. I really don't know the number these days. Quite a few studies.

Let me show you, again, if one is interested in just the spine, this is not an infant spine. It happens to be an adult image there. But one can localize and make a measurement just at the spine for bone. But for body composition work, we do total body measurements,

which is shown here. I think you can get an idea of the skeleton seen. I think you can see the soft tissue parts that are obviously not the bone, and we can then get that information.

Now, as you heard earlier, the BMI in terms of the height of heads, if one chooses to, one can actually decapitate the image and just worry about this part if you're concerned about how does the head contribute to all this information. So it is possible to do that kind of stuff.

By the way, the time it takes us today--in 1988, if you attempted an infant, which you would never be successful at, it would be 25 minutes. Today we do an infant in less than 3 minutes with the newer scanners.

So DXA, what are the advantages and some of the disadvantages? First, DXA has almost achieved a reference status within the body composition field. It still has some improvements to be made, but, again, in terms of everything else, it is the better technique that we have.

Advantages, it does give us good precision and accuracy. It is the only technique for a single assay that gives us basically a 3-compartment model: bone, fat, and lean. As you can see from that image, we can get some regional information if we choose to do that. It has a very low exposure risk. There's a very minimal

amount of that. And there are more the reference populations out there for adults, for children, and they are being developed, and several for infants, if you know the references to look for.

The disadvantages, very low exposure risk, the same thing. One could argue whether it's an advantage or disadvantage. One of the ways I talk about this--in fact, I just thought about it coming over here yesterday on the plane--was on the flight I was on, there was at least five children under the age of 2 on that flight with me, and the radiation dose they got on that flight exceeds what you get from the DXA. In fact, it's two to three times higher. So there's an idea what the risks are involved.

Scanners are not optimized for infants. If you get a scanner, it's adult size. They have not--the industry has tended to resist this, primarily because of the market that they are focused at, which is osteoporosis in older women.

It doesn't give us a 3-D image. You saw that 2-dimensional image. It gives us 2-D not 3-D imaging. That's what I call boot-strapped 3-C model. It's not a perfect 3-C model, but it's not bad. And the one different problem has been that the results differ between manufacturers, so that if you do a study--a

multi-site center study, you want to stay with the same instrument, same software.

Okay. Precisions and accuracy of different body composition measurements and the minimal detectable change in an infant. For this I chose basically a full-term infant and made the assumption it's 15 percent body fat.

If you look at these methods, the water, the dilution method, or the bioelectrical impedance method or TOBEC, this is the potassium one here. DXA, the bottom three, the fat, fat-free, and the bone, precisions, these are optimistic. Precisions tend always to be a little bit better than--whenever you do a precision measurement study, they always do much better than they do random. I'll guarantee that.

This shows you the precision measurements here. This is generally the accuracy. Precisions are, let's say, in the 1 to 5 percent range or 2 to 5 percent range. Accuracies tend to be in a 3 to 5 percent range.

If you take these, this information, and take this size of an infant and you translate those into one of the minimum detectable changes for that infant, the values are shown here, the last column on the right. And the percents are those percents of what that person had in terms of the composition at that age.

So one can measure water, changes at 5 will start to show up. If you use (?)-ium, they can get worse, TOBEC or BIAs. Fat-free mass, 125 gram changes, only about 5 percent of the total fat-free mass. Forty grams of fat, if that starts to change, it's 8 percent of that to implement this weight and composition. You can start to see changes relatively quick with the single input.

This shows the relationship between precision of the methods and what kind of a change, minimal change is required for that to become statistically significant, at the 5 percent level and a power of 0.8.

I call this the clinical application in individuals, what I consider to be a clinical application, what's changed in that individual. And if you look, for example, this is the relationship that the minimal detectable change--this is approximately 3 times the precision. And so that if our precisions for BMC are somewhere around 2.5 percent, 6 percent change would occur, lean tissues at about 10 percent, and fat-free mass at around--if the precision is 4.5 to 5 percent, we'd have to see changes in the range of about 14 percent to be significant for that individual.

Now, this I show you because this is the difference between 1988 and today. I only got 30 seconds left? Oh, well. Well, okay. Very quickly, this shows

you the methods. This shows you what the precisions are for FFM, fat-free mass. If you translate those into fat-free mass, this is the tail here. The top three were in 1988. That's why you couldn't do it. DXA is hugely improved since then. This shows you the kind of weight gains that would have to occur in infants at, I guess, again, the standard term infant here, very small weight change with DXA, large weight changes for--I'm out of time. She says zero. Anyway, the--I'll keep on going anyway. And the number of weeks would have to be changed in that individual.

I can't believe I've taken up all the time already.

This just shows you the rates of change that occur with age. You can find this from several different sources.

This shows you the first six months where you have to figure out what the rates of change are in the composition, and here's a series of papers that are on infants. There's one on TOBEC. This one used a series of methods which we can do. All the rest you'll notice, with the exception of one here and dilution, were done with DXA. The weight ranges are shown here. The number of infants are shown over here. We are now doing a meta-analysis to bring this together into one common reference database.

And I want to just quickly go through these, again, precisions, 2 percent, 3 percent, 6 percent reported here. And if you did a calculation--this is an interesting study in twins. They looked at the weight difference, which is about 14 percent, and they calculated again with these alpha-5, power of 0.8, 40 to 45 infants would be needed to detect a 15 percent difference in one or more of any of these three compartments.

Another paper here, let's see, this particular paper does give percentile curves for each of these values as a function of weight. And, again, what I'll point out here is that even though you may have a 3.5 kilogram weight infant, the fat range can range from 10 to 26 percent. By the time they're 10.5 kilograms of weight, it can range from 22 to 23 percent. So weight does not represent fat.

This paper is another one that has percentile curves of each of these compartments versus weight. Again, comparable to the other one, actually these are lower fats. Interesting, this is an European study. The previous one was a U.S. study.

And I want to show you--this is the last one, slide here, and I'm a minus five I think now. But here this shows the changes--total body DXA. These were preterm infants. Initial weights were about 17--under

1750. This was fortified human milk formula, and this was a preterm infant formula. And the baselines are measured at 3 weeks of age, repeated again at approximately 3 to 4 weeks later. And this shows the statistical ability to measure changes in body weight, for example, and in the fortified human milk you can make--you can change--see the difference at this level. And the preterm formula, it's about 19.9 grams per kilogram per day. Differences were four, and this is statistically different. You can see these--in other words, you can compare in groups of 20 versus 30, you can see differences in weight, you can see differences in lean mass and fat mass and bone mass by DXA at three weeks between these two groups of children. So it is possible with DXA to measure not only weight but the composition of that weight change in relatively small sample sizes.

And if you want to convert those to growth kind of numbers, this represents about 2.3 grams per kilogram per day in terms of growth, and that's about 12 percent of the mean weight gain in terms of composition. This again is 2.1 grams of lean mass per kilogram per day, 15 percent, about 1.2 grams per kilogram for fat mass and 76 milligrams per kilogram per day for the BMC.

So the point here is that we can measure--I think there was a question, could we measure changes of

composition that would be comparable to 3 grams per day? The answer is--from this study the answer would be yes, we could do that. Again, relatively small sample sizes of 25 to 30 children.

I'll end there. Thank you.

DR. GARZA: Thank you very much.

Any questions or comments?

DR. MOYER-MILEUR: I have a question. When you do your measurements in your babies, are they sedated? Because we find that we require sometimes more than two technicians to keep a baby quiet to minimize the movement artifact.

And my other question is, with preterm babies, we found it somewhat difficult to do early measurements because of equipment artifact, that they have leads and monitors on that make it very difficult to get a true assessment using DXA.

DR. ELLIS: Yes, two things. One, none of the infants that we measured and none of the infants that any of these studied were sedated. These are--again, they're all healthy children.

Our experience has been if you feed them right before you want to do the measurements, they tend to be rocked in the chair by the mother or someone, they go out, and then you can make the measurement pretty easily.

I don't know which machine you were using. Was it a 4500A or 2000 or--

DR. MOYER-MILEUR: We have a 4500A, and we also have--

DR. ELLIS: You should do it in three minutes or less. You do have to work at it. I mean, normally with older children who will cooperate, that could be anywhere from age 5 to 18, depending upon what you're looking for. But those children can get on a bed and will cooperate and can be--and they'll do it.

Here you can do the whole procedure in ten minutes or less. Here sometimes you have to spend as much as an hour to get the one measurement done. You have to work at it. But none of these children were sedated. We don't sedate any children in any of our studies.

DR. MOYER-MILEUR: Yes, and I just, you know, would caution DXA in the infants in that it requires people with specialized training so that you can't just--

DR. ELLIS: Yes, yes.

DR. MOYER-MILEUR: --go to a community hospital and get their--

DR. ELLIS: Yes, you can't--you cannot send these children to a radiology department even with the hospital because they simply are not experienced with measuring children. They just don't like it when

children show up. They don't have--infants, they basically will send them back. They will not--they will not take that hour, hour and a half to do it in. It takes effort sometimes. Sometimes they go right on the bed and out. It's always the ones that show up at 4:30 that take the hour and a half to two hours, though.

As far as the artifacts, you're right. You have to be careful about artifacts. You can delete those off the images, though, pretty well. If you take leads out, for example, out to the side, you can delete those right off the images. And so that's a minor effect if you deal with it right.

DR. GARZA: Dr. Denne?

DR. DENNE: I was wondering if there are any direct comparisons in infants between skinfold thicknesses and DXA for fat mass.

DR. ELLIS: There may be a few, but, again, the issues have been that skinfolds are probably more difficult to get than the DXA. We have skinfolds in some of our kids, but we just don't rely upon them for anything.

DR. DENNE: It would be an interesting comparison to make. You know, relative difficulty depends on what you're actually used to doing.

The other question is: How is DXA validated in infants? I mean, most of this body composition, you

know, was validated against the other techniques which all have their own sets of issues.

DR. ELLIS: Yes. The validation of DXA are done two ways. One is with animals, small animals have been done. We have done 73 piglets under the weight of 10 kilograms. Other people have done comparable size piglets. And probably if you add everything up, it's probably about 200 pigs have been done over the years at different centers with different machines.

The pig is not the best of models because, for example, its bone is more mineralized than infant's. Weight-wise, composition-wise, soft tissue is not that bad.

The other way we've done it is we actually built phantoms. We've actually fabricated mock-ups of the human body with parts made from polyester resin, doped with calcium and phosphate--phosphorous compounds to simulate that. But that's how it's done.

I have also done cadaver work. The problem--not the problem. The situation is that, unlike Elsie Widdenson, today's environment would not allow one to chemically digest the infant body, so we have done that in about 30--more or less 30--these are all preterm infants, and we did that by a technique called neutron activation analysis where we do a nuclear, chemical--nuclear chemistry technique where you measure calcium,

phosphorous, sodium, chlorine, phosphorous, manganese and magnesium and potassium.

And so if I look at the BMC bone versus fat, and if I look at the other ones and make some model--I have to make some modeling assumptions now about how much sodium is in the water, extracellular water and so forth, but they come out pretty well, with the 5 percent kind of accuracies.

DR. THUREEN: In the past several years, it's been recommended that at different centers, even if you have the same type of machine, you should do your own phantoms. Do you think there's enough phantom data out there now that that doesn't need to be done? Or if you're going to do a multi-center study, do you think that needs to be done?

DR. ELLIS: Well, for the multi-center study, there should be at least a common phantom that is going around to all those sites. One, to do the initial calibration to be certain everybody is within reason of the numbers, and then continue on throughout going for the study. That's typically what we do in all studies at all ages, whether it's infants, children, or adults. That's what we do these days. For multi-site studies, there's a common set of phantoms that go around all the time.

DR. GARZA: Dr. Stallings?

DR. STALLINGS: I'm one of those other six people in the world that's got the TOBEC, so I agree those, you know, have tremendous advantages. But I don't think that we would be able to use them. So I think, you know, bringing us to issue with DXA and how we could use it is an important question.

Laurie asked one of the big questions that I'm always asked, which is about sedation, and I would agree that, you know, natural sleep and that sort of thing and working in the research setting.

The other question sometimes is: How many images do you really have to take to get the one right? I even noticed on your slide the hand is--

DR. ELLIS: The hand was a little off.

DR. STALLINGS: And I just spent last week working with DXA and trying to figure out which one had all the body parts there and minimal movement.

But would you share with the group, you know, how frequently do you need to do two scans or you get halfway through a scan and then you do it again to get a good research quality measurement?

DR. ELLIS: I would say it's definitely less than 10 percent that we have to repeat the scans. It's like you say, there's a technician there. The image is being acquired while the scan is being done, and you can stop it immediately to start again, as you well know.

We have some--we have looked at some scans where the infant has moved, but we finish the scan and then repeat the scan again and looked at those. It has a lot to do with what kind of movement you have. As you well know, you can--if the child's arm is here--or say here when it starts and here when it ends, you have a three-armed child in the image you end up with, because it was here the first time you scanned through and caught on the second, on the lower case.

We have found that if we have motion in this direction, there tends to be a minimal effect because you're not changing anything. You're just moving the slice over a little bit here. But it's when there's movement like this, a flapping of the arms or kicking of the legs, if they're doing that, we don't scan them. We stop. But that's usually what happens, they wake up or something. Less than 10 percent.

DR. GARZA: Any other questions?

[No response.]

DR. GARZA: I have two, Ken. How well described are the specs of equipment that one would need to be able to measure infants reliably? Is there pretty much a consensus on the quality of the equipment, the DXA equipment that would be needed?

DR. ELLIS: You'd have to have something that's equivalent to what's called--there's basically two

manufacturers in this country. One is Lunar, the other one's Hologic. You have to have at least the DPXL for the Lunar at least the 4500A or DelphiA for the Hologic.

We always use Hologic's, and, in fact, we, you know, are constantly trying to improve those machines. I'm not going to tell you they're perfect, but they're the best thing we have. I think they could make them better.

DR. GARZA: You also indicated that individuals or personnel had to be specialized or had to be highly trained. How much training do individuals need to be able to use this equipment reliably, or was the training in reference to just training and dealing with pediatric populations?

DR. ELLIS: It's more dealing with the pediatric populations. It's more of that than it is simply for this, because once they understand--basically what you want to do is you want to minimize motion and have them in the right position and things like this. But it's more dealing with the pediatric population, dealing with a child that may want to cry for 20 minutes or something, or a mother that could be apprehensive when she hears the child crying. It's more that issue than it is anything else.

DR. GARZA: Dr. Stallings?

DR. STALLINGS: I just want to ask a little bit what Bert was doing. If we were doing such a study, a multi-center study, what would your advice be about centralized reading of the scans, the technician at the instrument site?

DR. ELLIS: It is a good point. These days, again, it is common practice now to send all the scans to a common central reading site because at least what happens there--well, I'm thinking more of the adults. If there is any kind of bias--in adults you set regions of interest. In the infants, it's a total body scan. There's no region of interest set. So it's less of an issue there.

But, again, the judgment about good scans or bad scans would come from one source and not from different sources. So it would be a reasonable thing to do.

DR. GARZA: Any other questions?

[No response.]

DR. GARZA: Thank you very much, Ken.

DR. ELLIS: Sure.

DR. GARZA: Committee members and guest speakers are invited next door for coffee. We're going to break right now, instead of at 10:35 as on your schedule, so that we don't break up the following three presentations. But I will ask everyone to try to get back here at about 10:25, 15 minutes from now, so we can assure that we

don't eat into any discussion time with either the speakers or anyone else.

Everyone else that is not a speaker or on the committee is invited to the cafeteria. These are federal rules. I didn't make them.

[Laughter.]

[Recess.]

DR. GARZA: The committee is seated at the table. If I can have our guests please take your seats, we're ready to start.

Our next speaker is Dr. Frongillo. Dr. Frongillo is an associate professor in the Division of Nutritional Sciences at Cornell University, and he's going to give us an overview of the World Health Organization Growth Reference Study that was referred to a bit earlier by Dr. Chumlea.

DR. FRONGILLO: Good morning. I'm going to stand here in the middle and use this archaic technology. It might help if we could dim the lights up at the front here a little bit.

These are some growth data from a single child, and you can tell what country they're from if you look at the units of measurement in pounds. And if we plot the data, it's actually more interesting to look at. We see a trend like we expect. But to try to really discern anything about what the pattern is, it's helpful to

compare it to something. And so this is a graph--this is the old U.S. reference, the 1978 reference, and this is the same child. And you can see this graph on the left, the child started off at the bottom of the distribution and then seemed to climb into the chart a bit, and then at about four months or so started to really--three to four months in there, started to really--its trajectory is now falling well below the chart. And eventually by--this is about two years where you see it comes back onto about the same percentile where it originally started.

This discrepancy that we see is either telling us something about this particular child or it's telling us something about the reference.

Well, it turns out it's telling us something about the reference because that's a breast-fed child. And if we look at a data set, this is a comparison that the WHO infant growth--an analysis that was done in the early 1990s, and what it shows is that, if you look on the left here, this is for boys, this is weight in kilograms and age going up to 12 months. And the dotted line to the 1978 U.S. reference which was adopted about that time, just after that by WHO, is the international reference. And the breast-fed data set, these are infants who were exclusively breast-fed for four months and then continued breast feeding through the first year.

And you can see again that about four months you start to see the solid curves deviating from the dotted lines, which is the same pattern that we just saw. And so this is showing that in a sample of about 426, I think it was, infants epidemiologically what we just saw in that individual child.

And a similar pattern was seen for girls. We can amplify this in a way by--what I've done is just simply take the current reference, the current international reference, the 1978 reference, and that's what would be at zero. So if these children were growing exactly like the reference, there'd be a horizontal line right at zero here. But what we can see is that this breast-fed set seemed to grow, if anything, a little bit faster at first, and then by the end of the first year had increased their weight substantially less than the U.S. set.

In fact, if we calculated the difference in rates for these two groups from zero to 12 months, it's about 2.7 grams per day. So this is something to keep in mind for later when we're thinking about how big our meaningful difference is.

The difference from about one month, which is at the top there, from the maximum to the minimum at 11 months is about five grams per day.

So that was for weight. If we calculate a z-score--and we'll have--I'll put this up here because we're going to hear about z-scores at various points. A z-score is where we take a particular measurement for a child and compare it to a reference median, whatever the reference is, and then divide it by a reference standard deviation. So the graph I just showed you just showed the numerator there, but for a z-score we also divide by the standard deviation. And the reason we do that is because then it's easy to imagine that the growth of a, quote, normal population would fall between about minus two and plus two z-scores. About 95 percent of the distribution would fall there.

So if we look at z-scores for this breast-fed set, then what we saw was that, regardless of the index that was used, whether it was length-for-age, which are the triangles, which is the curve near the bottom, whether it was weight-for-age, which is the circles, or weight-for-length, which is the squares, we saw a very similar pattern with this breast-fed set in comparison to what was then the U.S. and international reference.

So this discrepancy, along with other information that was obtained during the review that the Infant Subcommittee made during the early 1990s leading up to the WHO publication in 1995 of the uses and interpretation of anthropometry, the recommendation was

made that consideration should be given to making a new international growth reference.

The justification for having an international reference is, first of all, that it allows cross-national comparisons to be made that otherwise couldn't be made; and since there's been an international reference since the late 1970s, it's allowed us to do some things in a comparative way globally that we weren't able to do before. For example, this is a graph from a WHO publication that was in the bulletin of the WHO showing that the trends that have occurred from 1980 to about now so that we could actually look at the progress in Africa on the left, the very rapid progress that occurred in Asia. This is in percentage--the percent of the population that's stunted--and the rapid progress that was made in Latin America and the Caribbean region.

This kind of comparison has been made possible because there is a common reference being used throughout the world.

In addition, we know that it's very expensive to make local references, and also that in developing countries where there's still a very strong cyclical trend in growth, if a local reference were made, it would have to be revised very quickly because of changes that are occurring.

The justification for having an international reference goes back to work that was done in the early 1970s. This is a well-known graph from a well-known paper, in '74 I think it was, showing that these curves right here were all curves of high SES children in well-off countries, whereas the ones that were down here were children in developing countries that were not so well off. And so the fact that these were all so close together meant that growth roughly from one place to another where children are growing in conditions that are favorable to growth tends to be roughly about the same.

Some work that the WHO has done with a cross-national data set collected by the Human Reproductive Program recently shows that--this is for girls--across a number of different countries, these are children who were reasonably well off SES, not necessarily the highest SES, showed basically that, with the exception of this lower curve, which is in China, that these other curves all pretty much are very close together. Again, giving more recent evidence of the idea that it was reasonable to make an international reference where data from multiple countries could be combined.

So let me tell you, then, a little about the effort that's underway in the Multi-center Growth Reference Study.

First of all, I wanted to point out that a reference--the idea of a reference is that it's a tool for providing a common basis for the purposes of comparison. So we're interested in references because it allows us to compare as opposed to a standard which then involves a judgment. So here we're talking about making a reference, and during the early 1990s, both the U.S., in preparation for the revision of the U.S. reference, and also WHO examined the current reference which was being used in both the U.S. and internationally. And the sample that had been used for the early infancy especially was from Fels, which was one particular place in Ohio. The measurements were taken every three months, and in the very early period we might wish for more than that.

There were very few infants that were breast-fed for an extended period of time, and at the time that this reference was made, there simply wasn't the technology to do curve-fitting that we now have.

So those reasons, plus the main factor that the breast-fed infants seemed to grow differently than infants who were not necessarily breast-fed according to feeding recommendations, drove the decision to make a new reference.

At the time the WHO feeding recommendation was that infants should be breast-fed exclusively from birth

up to about four to six months, and then after that they should continue to be breast-fed for up to two years or beyond. You may know that WHO recently--I guess about a year and a half ago--revised this to be from birth to about six months for exclusive breast feeding.

So the objective of the Multi-center Growth Reference Study that WHO is doing is to build a set of growth curves for all children under age 5 years to be adopted as a new international reference for assessing the growth and nutritional status at both the population and individual level.

When this effort was started, it was clear that there were a couple of conceptual issues that needed to be thought through. One was that some references have been constructed, especially, for example, in the U.S. or in England, to take two examples, have been constructed to be descriptive references, meaning that they were intended to describe the growth of the population at a particular time.

This is different than what's going on in the Growth Reference Study, the WHO Multi-center Growth Reference Study, which we can think of as perhaps a prescriptive reference, meaning that it's meant to be a reference that depicts the growth of infants who were fed according to current recommendations for how children should be cared for during infancy.

The other issue had to do with maximal growth versus optimal growth. In the past, we've had a tendency to think that maximal growth and optimal growth are the same thing. The graphs I just showed you indicate that when infants are breast-fed, at least during the first year, and perhaps into part of the second year, they are not the maximum size they would be if they were breast-fed, but we think that because they're fed following feeding recommendations that that corresponds to optimal growth.

So the design of the study involves multiple geographically diverse sites. There's a longitudinal component which goes from zero to 24 months. And each site was asked to recruit about 300 infants per site in the hopes that at least 70 would be available for inclusion in the final reference. We've actually done better than that because the compliance with the feeding recommendations by the mothers and infants has been much higher than the 25 percent that we feared might be there. So we've actually ended up with quite a bit more than 70 percent. And then very frequent measurements, I'll show in a minute, and then there's a cross-sectional component which overlaps the longitudinal component. It starts at 18 months, goes up to 71 months, past 5 years, to make sure we have enough data on the right-hand side to be able to characterize growth well up to at least 5 years.

And sites were asked to recruit about 1,400 per site, which, again, would give a minimum at each age of about 70.

In the longitudinal component, which is a very demanding part of this study, measures of weight, length, and head circumference are collected frequently during the time. At birth, there's one visit, of course, and then in months 1 to 2 they're biweekly. So there's four visits there. In months 3 to 12, measurements are monthly, so there's 10 visits for that. And then in roughly the second year, they're bimonthly, which is six visits during that time. And then arm circumference and skinfold measurements are also taken in the same schedule, starting at 3 months, as in the other measurements.

Now, the way the study was constructed, there were a set of criteria at the population level and then a set of criteria at the individual level. So at the population level, the idea was to find populations of infants who did not have socioeconomic constraints on growth, where mobility would be low so that they could be followed, where at least 20 percent were willing and able to follow the WHO feedings recommendations, with support, so there had to be existence of or at least the ability to build breast-feeding support systems; and then there had to be local presence of collaborative institutions

who were capable of carrying out this kind of exacting work.

Then at the individual level, individual criteria were set that there was an absence of health, environmental, economic constraints on growth; the mother was willing to follow the WHO feeding recommendations; that the mother was a non-smoker; that gestational age would be at term, which we defined to be 37 to 42 weeks; and that the infants wouldn't have any severe illnesses that would be expected to affect growth.

The protocol for the study site selection then applied to subpopulations the fact that socioeconomic status did not constrain growth, it was low altitude, low mobility, the minimum of the 20 percent, existence of breast-feeding support systems, the local institutions. We looked at the rate of hospital deliveries because we had to know that there were enough infants being produced quickly enough that they could be enrolled in the study so we could get the study done sometime in our lifetime; that there would be sufficient numbers of eligible birth; and that it was feasible within those locations.

In some places, for example, in a really huge city of 10 million, it's just not feasible to do a study like this. The logistics are too difficult.

The Steering Committee also considered some other factors in its thinking about in looking at mean

birth weights, maternal heights, complementary feeding practices, health-related behaviors, and the existence of environmental hazards. The Steering Committee looked at geographic distribution. It's a global reference, so WHO tried very hard to have geographic representation throughout the world, and funding issues, because it's expensive to do this kind of study and we had to think about where the funding would come from and how that could be arranged.

The protocol was developed by this set of characters here. The main reason I put it up here is to show you that it's a multidisciplinary set of people representing a diversity of backgrounds and institutional relationships. So this group put together the protocol roughly in the '95, '96 time frame. And then the study is being run now with an advisory group: Cameron Chumlea, Tim Cole, myself; Ray Martorell is the Chair of this group; John Van den Broeck, who recently moved to South Africa; senior scientists representing CDC and UNICEF, previously was Roger Shrimpton in UNICEF; and then WHO, the day-to-day work gets done at the sites and at WHO Secretariat. Mercedes de Onez coordinates all of this with her staff there. A very dedicated group of people.

The sites that have been selected are the following: in Pelotas, Brazil, in the south part of

Brazil near the coast; Victoria is the PI with Cora Post; in Oslo, Norway; in the U.S. at Davis; in Muscat, Oman; in Accra, Ghana, the capital of Ghana; and in New Delhi, India. These are the six sites. Each of these sites represents very differing and very large challenges to carrying out this study. And it couldn't be done without the commitment of the teams there who are doing the work on the ground.

The Steering Committee is chaired by Cutberto Garza representing UNU. Data management is done by local data entry and checking at the local, each site, and then the data are shipped to the WHO Human Reproductive Program. They have extensive experience in handling large, multi-country data sets, and they've done a fantastic job in coordinating all of this and in helping to ensure data quality.

The decisions and information about the study, there was a working group on the growth reference protocol. We have Steering Committee and Advisory Committee meetings periodically, other meetings. Particular tasks are handled by other meetings. We do a lot through electronic mail. Various site visits are made, were made before the study started in preparatory work and are made throughout the study, and rapid surveys were done at the beginning to get information that was

needed to actually do the planning for the data collection in the particular sites.

To give you an idea of what's involved, I just made a list here of sort of the documentation that's been produced, which will give you a feel for what was necessary to carry this off at the level of scientific quality that was being strived for. The protocol was developed, a measurement of standardization protocol was developed. A manual of operations, a generic manual of operations was produced, and each site had to adapt that for its particular location.

There was a protocol for the 12-month visit. A special effort was made throughout this study on the epidemiological quality of the study to follow all infants. Even if mothers wanted to drop out or weren't complying, an effort was made to follow them as much as possible so we could keep measurements for every infant even if they weren't exactly following the feeding recommendations.

In particular, we had an effort made that at 12 months we could go out and get at least some measurements on all those who were not willing to continue.

There were guidelines for complementary feeding. A protocol was developed for assessing diet, for the cross-sectional study, for data management, and then questionnaires were produced for both the longitudinal

and the cross-sectional study. And if any of you have been involved in questionnaire production, you can imagine what those meetings were like as people argued about the exact wording of every question and every answer.

So that's an overview of the study. Basically where we are now is that the last site--the sites have been selected over time. Some were able to start earlier than others. Some had technical challenges that took longer than others. So in the next few months we'll be finished with data collection. There's a meeting coming up to look at and try to decide on the final methodology for analysis. Some preliminary work is underway. So this will be proceeding, and we're talking about having a reference be available in the 2005 year.

There's a lot of work that needs to be done in order to prepare for not just producing the reference but preparing for how it will be depicted and how it will be used and testing that will have to be done with the reference under the auspices of WHO.

DR. GARZA: We have about five minutes left for questions. We might be able to go over.

Dr. Stallings?

DR. STALLINGS: Well, one, to compliment the group. This is an extraordinary effort and an extraordinary study. But to cut to the chase, do you

think that we will see this used in the U.S. as the growth standard for infancy through 5 years or even infancy through 3 years with what we think of now as our traditional infant chart? So I'd be interested in your opinion and certainly in relation to what we're here for, which is to start to understand the best comparison group for children in the U.S. who are taking infant formula.

DR. FRONGILLO: Okay. Well, I think I'll probably not try to answer your question so directly. Let me just say that I think the advantages of this reference will be that it's longitudinal, that the longitudinal data in the first couple of years will have the ability to look at--to have a reference, a velocity reference, and so a judgment will have to be made whether that's better than, for example, the Iowa/Fels data that are available.

The second thing is that we know that infants who are fed following the breast-feeding recommendations will show a different pattern of growth than formula-fed infants. So to the extent to which it's seen as desirable to have a reference which fits that growth pattern for infants who are being breast-fed, then I think that would certainly be the advantage of the new reference.

DR. GARZA: It may be useful to describe the references that are going to be available. Is it just weight and length, or are there others?

DR. FRONGILLO: Well, those reference data will be available for all of the measures that I showed, so we will have data on weight and length, head circumference, arm circumference, and skinfolds.

DR. GARZA: Any other questions?

DR. THUREEN: Why did you choose to include skinfold measurements, and what kind of information did you hope to get from that?

DR. FRONGILLO: Well, I think that there was a debate about, you know, the importance of this and when it should be started, and I think the lack of reference data on skinfolds was very compelling and it was thought that, given the potential usefulness of that information in the future, that while a study of this effort was-- while this large effort was being made, it would be important to have that kind of information available.

DR. GARZA: Dr. Baker?

DR. BAKER: I have a question about the prescriptive nature of this. If you're going to do a study like this using a prescription, it assumes the prescription is right. It also assumes that it also would change, presumably, if the prescription changed.

Now, the WHO has changed it since this study was done. Does that make a difference?

DR. FRONGILLO: Well, certainly we thought about this a lot during the time in which the study was being planned. I don't think anyone imagined that the basic recommendation about breast feeding and complementary feeding is going to change appreciably, and at least not for quite some time.

Now, if it does, one of the things we wanted to do and part of the reason for the intensive follow-up even for infants who didn't exactly follow the current feeding recommendation was that it allows us to have the information available so that if 15 years from now we have new knowledge and decide that some slight revision of the feeding recommendation is made, anybody who's ever thought about changing the feeding recommendation will run after that possibility.

But, anyway, if anybody gets brave enough to try to do that, we will have the information available from the cohorts in the six sites so that one could conceivably reconstruct the reference to conform with that.

DR. GARZA: Yes?

DR. DOWNER: Have you decided exactly which tool you will be using to do the skinfold measures? And my second question is: Because what we consider SES for

different world populations differ so widely, how are you going to decide on what to use in this study?

DR. FRONGILLO: Okay. The skinfolds, which tool, do you mean which skinfold caliper? Basically I think--and Cameron can comment on this because he's the expert here. But my experience is that it's not the caliper that makes much--any difference at all in the measurement. It's the quality of the enumerator and their training in using the instrument. So that's not going to make any difference.

DR. DOWNER: What instrument have you planned on using?

DR. FRONGILLO: The instrument we are using is--what is the instrument we're using? Holtain, right. Okay.

And the second question? I'm sorry.

DR. DOWNER: The SES.

DR. FRONGILLO: Oh, the SES. In each site, a survey was done before the study began, the main study began, to actually look at the relationship between socioeconomic status and growth so that we could develop in each site exactly what the criteria needed to be from a socioeconomic standpoint in order to ensure that the population of infants selected was at high enough SES to not constrain growth. So that was done separately in each of the sites, and some of us traveled around to

different sites to help them actually carry that out. And you're right, in each site different criteria were needed because the conditions were different.

DR. GARZA: Any other questions?

[No response.]

DR. GARZA: All right. Thank you very much.

We'll move on then to Dr. Larry Grummer-Strawn, who is the branch chief of the Maternal and Child Nutrition Branch at CDC, and he will tell us about the NCHS/CDC's growth charts.

DR. GRUMMER-STRAWN: Good morning. If we can figure out how to forward this? Which one? Just here? Okay. Thank you.

This morning I want to give kind of an overview of the new growth charts and contrast them to the old NCHS growth charts. I'm going to start off with kind of a historical perspective. I'm sure many of you know the history but to kind of just set a context for all of us, do some comparisons of differences and similarities between the old and the new, and then go into some of the differences a little bit more explicitly, and then finally end with some analytic issues that the charts pose for us.

The original NCHS charts were released in 1977. Those charts were only percentile curves. They were

published by Hamill, et al., really became the standard of reference for all U.S. infants.

Subsequent to that, there was a normalization of those curves at CDC. Those were actually published in 1987, but were actually available for use long before that. So people who were interested in normalized curves had access to them earlier. And, finally, the WHO adopted those curves as being the international reference, really referring to the normalized curves. The adoption by WHO actually came prior to the publication of the normalized curves.

The reason that I point this out is that those curves never became one and the same. The percentile curves never matched with the normalized curves, and so someone who was using clinical charts that actually saw the graphs in front of them was not necessarily using the same cut-off points as someone who was using computer software or might be analyzing data sets. They were very similar to one another. They were analyzed off of the same data, but were slightly different from each other. And then, finally, in May of 2000, CDC released a revision to these charts.

Now, at the time that the original charts were created, NCHS was a separate agency, and so they were referred to as the NCHS Growth Reference. When people talked about the normalized curves and put them in an

international context, they might have referred to the NCHS/CDC/WHO because of the separate role each of those agencies played.

Subsequent to 1977, NCHS was actually incorporated into CDC, and so the new charts are referred to as CDC charts. That does not mean that NCHS was not an active player. They actually were the progenitors that moved the new charts forward. But it was in a larger context of CDC, and other parts of CDC were also involved.

So what are some of the similarities? First of all, both sets of charts are looking at the same indicators. We have weight-for-age, length-for-age, weight-for-length and head circumference-for age. Both sets are sex-specific. In neither case did we have any separation according to the parental anthropometry, race, ethnicity, infant feeding mode, different things that might impact on the growth of the infants.

Ed just described for us kind of the difference between the idea of a reference and a standard or a descriptive reference and a prescriptive reference. These clearly are references, not standards. The only kind of movement toward a standard is that very low birth weight infants, that is, less than 1,500 grams, were not included in the new charts, the CDC 2000 charts.

These charts reflect attained size, not incremental growth, and in both cases, we have accessibility of percentile scores and z-scores. In the new charts, the z-scores are one and the same.

The differences between the old and the new are, first of all, that in the 2000 charts, the data for infants are now nationally representative. These represent a broader spectrum of race and ethnicity across the United States, a broader spectrum of socioeconomic status, and there's an increased representation of breast feeding in the charts. That doesn't mean that they are primarily breast-fed children, but there's a mixture of formula-fed and breast-fed children.

The 2000 charts are based on a pooling of several data sets coming together, whereas the 1977 NCHS charts were all based on the Fels data set.

There were some minor changes to the smoothing techniques, which I will mention briefly. As I said, the z-scores now are one and the same as the percentiles. There's a one-to-one match on those, so it doesn't matter whether you use computer software or you're using printed charts. You're going to be looking at the exact same cutoffs.

Another minor difference is that the length now extends down to 45 centimeters rather than 49 centimeters when we're looking at weight-for-length, and when ow have

on the clinical charts the accessibility of 3rd and 97th centiles. Of course, based on normalized curves, you can get any kind of centile that you're interested in, but the difference is that the clinical charts that are produced actually do extend out to the 3rd and 97th percentiles, and the smoothing-out rhythm is intentionally extended out to those centiles to make sure that we're incorporating the original data out that far. And, finally, as I mentioned before, very low birth weight infants are excluded.

So what are the data sources? Well, in 1977, I'm sure you're all familiar with the Fels Research Institute study. It was done in Yellow Springs, Ohio, primarily represented Caucasian, middle-class families. And while not exclusively formula-fed, this group is considered to be almost all formula-fed infants.

There was a longitudinal follow-up study. The children were followed from birth, 1 month, 3 months, 6 months, and at three-month intervals after that, but we're focusing on the first 6 months today. The data were collected between 1929 and 1975, and there were a total of 867 infants.

Now, the data for the CDC 2000 curves, as I said, represents a number of different data sources. What I've put up here is a graphic showing you how at different ages, different data sets come in. I'm going

to talk about kind of these different data sets at different points.

The primary data source is the NHANES III data here when we're talking about children birth to 6 months of age.

Now, this is represented by the long line here in light blue. The NHANES III was the only data set for which we had nationally representative data prior to 6 months of age. Starting at 6 months, the NHANES II data were also available, and starting at 12 months the NHANES I data were also available.

Now, the reason that these are important for us today is that because we're smoothing these curves across age, the influence of those older data sets does come in at 6 months of age and at 12 months of age. The smoothing is across all ages, and so the curves below 6 months are also affected by those other data.

However, the NHANES III data started at 2 months of age, and among 2-month-olds, it was a fairly small sample size. In order to extend these curves down to birth, we had to look to other data sets. And in each case, whether we're talking about head circumference-for-age, length-for-age, weight-for-age, weight-for-length, in each case we have to turn to different data sources.

I'm going to start in the middle here to talk about in the weight-for-age. In this case, it was fairly

straightforward to use a birth point coming from the national birth certificates. We had all of the birth certificates during the years that the children in these NHANES surveys had been born and had the birth weights available on all of those. So it's a huge sample size, a very precise point that is truly not only nationally representative but a census of all births in the United States. And so in creating the curves, we were able to connect that particular point, actually anchored the curves to that point, and then smoothed it with the data starting at 2 months of age from the NHANES.

However, we don't have national data on any of the other indicators, head circumference or length. With regard to length, we did find that there were two states that routinely collect length data at birth in a representative fashion. We analyzed the data in those states for their birth weights against the national birth weight distribution and found that they were quite representative--those states were Missouri and Wisconsin--and felt that because the birth weight distribution matched the national distribution, we could expect that they should represent the national birth length distribution, even though we do not have data on the national birth length.

So when we looked at the weight-for-length curves, those are based on connecting the dots between

the Missouri and Wisconsin data on weight-for-length versus the NHANES III data weight-for-length, and those curves were connected together.

Initially, we intended to do the exact same thing with length-for-age, use only the data from Missouri and Wisconsin to connect these curves across age and with the NHANES III data. We did that in our first pass but analyzed--as we were evaluating the curves against alternative data sets, we found that we were comparing against the Chicago data set, the WHO pooled data set that Ed just described for us, as well as some of CDC's surveillance data, and found a common pattern in all three of those that the curvature between birth and 6 months did not match what we found in external data sets and felt that this was partly an artifact of the fact that we only had about 35 infants from the NHANES III data that clearly were not matching the normal pattern of growth. And so the curves were being pulled in the direction of those NHANES III data from a very small sample size.

So what we opted to do was to choose an additional data set to add in here between just beyond birth--these were not birth points, but at the first visit to a clinic--up through 5 months of age from the CDC's Pediatric Nutrition Surveillance System.

These are data on low-income infants. However, we didn't use a representative sample of low-income infants. Instead, what we did is we chose clinics that matched the national distribution in terms of their mean, standard deviation, and skewness at each age from birth--from 3 months of age through 11 months of age compared to the NHANES III data. So we were pulling out clinics that the children in that clinic happened to look exactly like the national distribution and chose those clinics and assumed that they would also look like the national distribution would have looked between birth and 3 months, and then added those data to the curves here and connected using the Missouri and Wisconsin data, the CDC nutrition surveillance data, as well as the NHANES III, and continued the curves using that.

Finally, for head circumference, we had no national data, and we returned once again to the Fels Institute data for the head circumference at birth point and, again, connected that with the NHANES III data.

So you see the picture here is one of bringing together multiple data sets. We had a number of comparisons to try and make sure that these were valid comparisons to make, but it certainly leaves us with a difficulty using multiple data.

So what are some of the other differences? First, with regard to the smoothing, the old curves were

smoothed with cubic splines, with knots at birth, 6 months, and 18 months, for those of you who work on these statistical arenas. What that meant for us was that there were six independent parameters that characterized growth between birth and 3 years of age.

In the CDC 2000, a completely different approach to smoothing was applied. Here we used fractional polynomials that had been used previously in other growth studies, primarily used in the Fels data as well as some Canadian data. And so they had kind of been proven methodologies for assessing growth during the first three years of life. However, there were a smaller number of parameters that described growth, really just three independent parameters--roughly three, because there were some other things that were done to the curves to get them to fit.

Finally, on weight-for-length, there were no set models, and so we used a 5th degree polynomial to maximize the flexibility of the curves there.

I mentioned before that in 1977 the standard deviations were estimated as a separate path and created a separate set of curves. In that case, we had two different standard deviations. There was a standard deviation above the median and a standard deviation below the median. And so if you think about kind of plotting the standard deviation as it goes across the curve, it

was a low standard deviation that instantaneously at the median rose to a higher level.

We calculated these scores in the normal way, taking the measure minus the median divided by the standard deviation.

In the CDC 2000, this kind of immediate change in the standard deviation at the median was thought to not be desirable, and so instead we had a more continuous change in the standard deviation. In this case, we transformed the data with a Box-Cox transformation, a power transformation. People understand how you take the log transformation of a data set or a square root transformation. The Box-Cox is a continuous set of transformations that you can then choose a parameter to say how much you want to transform that data to make it symmetrical. And then once it is symmetrical, you can fit parameters to normalize that curve.

Some other differences, clearly this group would be interested in the differences between the infant feeding in the groups. As I mentioned before, the old charts were virtually all formula-fed infants. The quality of the infant formula across that time, of course, has changed, so it is a mixture of a number of different kinds of feeding across those year '29 through '75.

In the CDC 2000, we have a mixture of breast feeding and formula feeding, but it still is primarily a formula-fed group. At 2 months of age, about half were formula-fed, half breast-fed, but by 6 months of age that was down to 28 percent currently being breast-fed. And we're not talking about exclusive breast feeding. As you can see, the exclusive breast-feeding rates are much lower than that. Down to less than 10 percent by 6 months of age were exclusively breast-fed.

If I can just take a couple of minutes to show you a comparison between the old and the new curves with regard to breast feeding, and we'll stop there, Ed showed a number of things as to how the old curves performed against WHO pooled breast-feeding data set. We did some additional comparisons seeing whether the new curves have actually improved that situation. So we've compared the WHO pooled data set that came together from six different studies of exclusively breast-fed children and pooled those data together. And instead of plotting the means, which is what Ed was showing us, here I'm going to show the percent below the 10th percentile using the old curves versus the new curves.

When we're looking at weight-for-age, you can see that there really has not been much of a change between the old curves and the new curves. We have this same problem that as children get older, we're going to

diagnose more of them as being underweight. Just as Ed was showing us that the means go down, the percent that would be low is going to get higher. And we still have that problem with the new curves, that it is considerably--we would have considerably more older children being considered underweight in this breast-fed data set compared to the younger children.

However, with regard to height for age, we've improved the situation somewhat. With the newer curves, there still is a tendency toward increasing the percentage that would be considered low as we get to older infants, but it's not as steep a trend as we had with the older curves.

And, finally, with regard to weight-for-height, whereas the old curves showed that same pattern of worsening nutritional status over the first year of life, on weight-for-height the new curves have pretty much wiped out that problem. We see a much flatter distribution across age.

I'm going to skip over these because Ed is going to come back to many of these points in his description of the analytic issues.

Conclusion: The interpretation of the new charts is really not widely different from the old charts. We're still using the same kind of way of

thinking about growth and the way we analyze growth as very similar in the old curves to the new curves.

There are a number of enhancements that argue for changing over to the new curves. I'm not arguing that we haven't made enough of a difference to adopt these new curves. However, I do think that the WHO reference that Ed has described would relate to a more substantive change in our interpretation of growth parameters, and we really need to give much more thought to different ways of thinking about growth than we have so far.

Thank you.

DR. GARZA: Thank you very much, Larry.

Any questions? Dr. Stallings?

DR. STALLINGS: I actually have a series of questions, lots of things jotted down. One, also, thank you and your team for doing this. It was a monumental change to give us this from the clinical point of view. It's wonderful to have the charts revised.

In that last set of slides where you were showing the less than 10th percentile, just so I'm thinking about it correctly, the perfect outcome would be 10 percent would be less than the 10th percentile, because if we were looking at a population study, would you by definition expect 10 percent to be less than the 10th percentile?

DR. GRUMMER-STRAWN: Yes. What I'm more concerned about is the pattern of growth there than the actual level. We're comparing curves that are based on formula-fed infants mixed with some breast-fed infants against a group of breast-fed infants. And so I wouldn't have been surprised if the level was somewhat different in a group of breast-fed children that might have less malnutrition than a general U.S. population. But I would expect the pattern should be representative of the pattern of growth.

DR. STALLINGS: But, still, when I was looking at it, when there was almost zero less than the 10th percentile, I saw that as unusual as when we have 20 percent less than the 10th percentile, that that's sort of the concept. If we were getting--whatever right is, but if we were getting it right, the population would go along that.

While we have all the experts in the room, we've heard three different people say three different things about whether the Fels data included a lot of breast feeding or not. And before you guys leave today or tomorrow, I'd like that to be readdressed. You know, you were very helpful in showing in your slide the percentage of children who were breast-fed and then exclusive, and so obviously it's not one number. Where you are at 2 weeks of age is different from 6 months. But I think it

would be helpful for the committee and for the FDA to have an understanding of what we believe the Fels data represented as exposure to breast feeding and what you believe your 2000 charts represent. This is--

DR. GRUMMER-STRAWN: I'd like to have that, too, because I certainly--

DR. STALLINGS: Okay. Well, then, you guys aren't leaving until we get this right.

[Laughter.]

DR. STALLINGS: The other part of that is I think getting a handle on that will be even more important as the WHO new charts come out and those data are there.

My last question, which really is a question, is: If I understand, then breast-fed babies grow faster during the first 4 months of life compared to our usual reference data and more slowly between 4 and 6 months, if we were to look at the zero to 6-month period that we're really supposed to be focusing on.

DR. GRUMMER-STRAWN: I think it's more like around 3. Is that about the peak? Three months of age. So faster in the first 3 months of life and slower after.

DR. STALLINGS: So that's part of the pattern that we're trying to capture that has made individuals difficult to look at.

Thank you.

DR. GARZA: Any other questions?

[No response.]

DR. GARZA: I have two. In thinking through the reasons why the WHO took more frequent weight measurements, one of the principal things that drove that was that very often the pattern of growth was used to assess the nutritional management of infants, the first 3 to 6 months.

To what degree can the present reference between used with that amount of detail, the first 3 to 6 months? Were the smoothing techniques in your judgment sufficient to capture the differences in growth patterns that Ginanne just described? Or did the smoothing eliminate much of that?

DR. GRUMMER-STRAWN: First of all, you have the whole difficulty of dealing with the cross-sectional data as opposed to longitudinal data.

DR. GARZA: I know. That was Part B to 1.

[Laughter.]

DR. GRUMMER-STRAWN: So all of those issues come in there. You have not a very large sample size in that age range, and so there is a fair amount of noise.

In addition to that, you're fitting basically a three-parameter model to the first three years of life. That doesn't give you a whole lot of degrees of freedom

to really let that first few months take on a particular shape.

That said, we did examine what the curves were doing in that age range, and they looked like they fit fairly well. But there's a large amount of noise in the cross-sectional data that bounces from month to month, and you look at the curves and say, you know, I think we did the best we can given the data that we have.

So I don't want to blame the three-parameter model, but I would say that if one had better data and wanted to capture really what is the pattern of growth in the first year of life, I probably would not do the smoothing in the way that it was done.

DR. GARZA: And the second, since you answered Part B of 1, we've been asked to look at various control groups, either historical or using specific references as controls in clinical studies. You also make the distinction between standards and reference. To what degree can, in fact, one use the present reference in making judgments, value judgments for clinical studies and making comparisons between feeding groups that might be placed on new formulas and the current CDC reference? Is it sufficiently robust to be used as a standard in making that clinical judgment that control groups normally play in clinical studies?

DR. GRUMMER-STRAWN: I think the question is one of what do you want to compare to. What is right? In comparing to these curves, you are implicitly saying I am comparing this child's growth or this group of formula-fed infants' growth to the way children have grown in the United States essentially over the last 10 years, maybe 15 years. Are you comfortable with that kind of a comparison?

You would say based on formula--in assessing a formula, you are saying this particular formula that we are evaluating generates a pattern of growth similar to the way children grow in the United States, whether they are formula-fed or breast-fed, fed on all kinds of different formulas, mixtures of solid feedings coming in at various ages. If you are comfortable with that kind of a comparison, this formula produces a pattern of growth like children in the United States, like a cross-section of all the children in the United States, then I say yes, this is a perfectly fine comparison to make.

If you want more of a prescriptive statement that this formula produces a pattern of growth that is the most healthy, I don't think that you can say that. I don't think that we can infer that a cross-section of infants from the United States with all of the variety of primarily infant feeding patterns--there are also varieties in terms of socioeconomic status, the kinds of

conditions children are running up against. I think that in terms of the impact of those on these curves, those are not as great. But the variation in feeding patterns, I don't think that we can say that we have the best pattern of growth here.

DR. GARZA: Any other questions?

[No response.]

DR. GARZA: Thank you again.

We'll move on to the next paper. I don't think Dr. Fomon needs an introduction. I can't think of a more senior person in pediatric nutrition than Dr. Fomon, and I'm very pleased that he was able to join us today and was willing to leave lovely Texas for Washington for this purpose, from one native Texan at least. And he will be addressing the Iowa data and the Iowa/Fels growth data.

Thank you very much for joining us, Dr. Fomon.

DR. FOMON: As the most senior member of the presenters, I'd like to make a statement for the presenters that we were asked to prepare a 20-minutes presentation, and if we run over and get the zero sign at 15 minutes, we're not very apologetic.

[Laughter.]

DR. FOMON: First I want to get out of the way what are the Iowa data and what are the Fels data, and then we can talk about more interesting things. The Iowa and the Iowa Fels data, Iowa Fels data is published by

Guo, et al. Term infants measurements all made with highly quality controlled efforts. Caucasian, we were able--and I will show you that--no, but it's in my paper--that the length and weight of the Fels and the Iowa series were very similar at three months, which gave us some encouragement in combining them.

Iowa data. There are other Iowa data, but I wanted to first speak about 8 to 112 days. These were all formula fed infants, 380 males, 340 females, and mostly infants of upper socio status, but not necessarily economic, because they were mostly UI personnel, University of Iowa faculty and students.

They were measured within 4 days of 8, 14, 28, 42, 56, 84 and 112 days. No exceptions. And the age of measurement by interpolation or extrapolation was made to the target age. So we adjusted if a baby was measured at 16 days. We used the 28, 16 and 8-day measurements to compute a 14-day measurement. It was fully longitudinal. There is no one of those 380 males or 340 females who was not measured at every time. This is published, so you can read about it in books.

Then we did also measure a number of infants from 112 to 196 days of age, and they were 165 males, 188 females, and there was a subsample of these that were also in the other group that I talked about. There were 63 males and 74 females who went from 8 to 196 days. We

have many more now, but it's too expensive to do the analysis.

The Fels data, there are a lot of Fels data. The Fels data that I'm talking about are the Fels data that are included in the Guo, et al. paper, and that included 240 males and 236 females, wide range of socioeconomic status, few measurements during the first 3 months. The target ages were 1, 3, 6, 9, 12, 18 and 24 months. Most of the children were measured within 3 weeks of the target ages, and there were some missed points, and they arrived at these ages by mathematic curve fitting. And there are other people who know a lot more about that here than I know.

So the Iowa Fels data, we ignored the Fels data during the first 3 months, so it doesn't matter whether they were breast fed or formula fed. There were, for the first 3 months, just Iowa data, 580 males and 562 females. The reason that's bigger than the numbers I showed you before was what I showed you before was formula fed. We included breast fed in this, and I'll tell you why.

Then we used both Iowa and Fels data, having established that the Iowa Fels size at 3 months was very similar. We used both for 3 to 6 months of age, and then we used only Fels data for 6 to 24 months of age.

So those are the Iowa data and the Iowa Fels data.

The Iowa data and the Iowa Fels data and most other referenced data, except those for international comparisons, have been developed to detect abnormalities of infants, of individual infants, and to detect abnormalities of individual infants you have greatest interest in the outlying centiles. The more individuals you have, the more confidence you have in those outlying centiles. What you want is early detection of growth abnormalities. Growth will not tell you whether a child is normal or not, but it gives the most important single clue to telling you that that baby is one that needs closer attention than the general garden variety baby.

And for that reason, weight gain is more important than length gain, and that's because it gives you this clue much earlier than change in length. I think I have that on the next slide. And you need data for at least the first 2 years. Iowa data are only good, up to at best, 196 days.

I have to go back, but I don't know how to do that, so I'll tell you that--can you go back? Just go on back.

The weight gain is more important than length gain because as I said, weight gain gives you the clue earlier. Length gain is very difficult to measure

accurately enough to be useful for determining changes in length the way it is done in hospitals, clinics and doctors' offices. So length is not really a very feasible way, and moreover, there are very few instances in which length gain will be abnormal and weight gain normal. So weight is the most important thing.

Now, when you evaluate an infant formula--we didn't think about this until much after we published most of the Iowa growth data--you have different criteria for what you need as reference data, and the characteristics are it should be longitudinal. It's difficult for me to agree that you can use cross-sectional data as a sensitive way of analyzing longitudinal data. If you're doing an infant formula study, you're examining how the infants grow over the period of study. For that you need longitudinal reference data. It should be gender specific. I haven't heard any argument about that.

The study integral must include at least part of the neonatal growth spurt, should include all of it or most of it. Neonatal growth spurt is from 8 to 42, maybe 8 to 112 days of age. After 112 days of age growth rates are substantially less, and we'll come back to that. And you need length data as well as weight data because it is possible that you would find--and I'll give you an example of this--a situation in which babies would grow

normally in weight or maybe super normally in weight, and the weight to length would be outside of what we see with usual infant formula, suggesting that maybe this formula is not fully adequate.

The reference population should be similar to the study core, and that's always going to be a problem. It will never be exactly like the study core, but the question is, how close can you get and how close do you need to get?

I missed the last one. If I knew how to use this, probably I'd know how to use my camera too.

[Laughter.]

DR. FOMON: The duration of study should be at least 84 days. That's a new minimum length. I made up the old 3 months minimum length that's in the AAP report, and I didn't have any good basis for that, but I think that maybe you can agree to 84 days as well as you can agree to some slightly longer figure, and we'll come back to it.

I said this. The most sensitive evaluation of the longitudinal growth study of a cohort in the longitudinal growth study requires longitudinal reference data. I don't say that this is gospel. It's just what I believe. Gains in weight and length are more rapid in infant males than in infant females. The formula may be

adequate for females but not for males. Nobody's arguing about the gender anyway.

The study integral must include at least part of the neonatal growth spurt. A formula may be adequate for older infants but not for younger infants. The reason is that during the period of most rapid growth the ratio of specific nutrient to energy is highest, and if you get beyond that period, the ratio of protein or calcium or whatever to energy may be down at a lower level. So if you start a study at 4 months and you get 4 or 5 months of additional data, the formula may be fine, but it doesn't tell you that it will be fine starting at 8 days or at birth.

Here is an example. These are males, and this is weight gain from 8 to 56 days, and we did a study of a relatively low protein diet based on isolated soy protein with or without a methionine supplement. We were interested in getting a fix on the requirements for sulphur containing amino acids. And with the methionine supplement the gain was 42.3 grams per day, and with no methionine supplement it was 38.8 grams per day. Compared to the reference data, 8 to 56 days, this was not significant and this was significant. From 56 to 112 days there was no difference.

I wish I had 3 or 4 more studies to demonstrate this, but this suggests at least that you need to have an

early portion where you have the maximum postnatal growth included in your evaluation of a formula. You need data on length as well as weight, and this was the best example. I think if I spent more time I could find more examples and then it would be more convincing, but I couldn't do that, because I had to work in my yard.

[Laughter.]

DR. FOMON: These are males and this is the BMI, and this is a low-protein formula and this is the reference. And at 8 days, when we enrolled them, the low-protein, the cohort receiving the low-protein formula had a BMI of 13. That was significantly less than the reference data. At 112 days the low-protein cohort had a BMI of 18.6 which was significantly greater than the reference data. And all our data, including--if you pardon my expression--skinfolds, indicated that these babies were fat. And what we speculate is that they were--and they took more volume. We, in all our studies, record how much the babies eat. We weigh the bottles in and the bottles out, and they took more energy in and they gained more weight, and we speculated that they ate more because of the low-protein content. They were eating--I don't know how they knew how to do it--but they were eating more to make up for the low-protein concentration in the diet. They got enough protein. They grew normally in length.

The study cohort must be similar to the reference cohort. They should be healthy, that is they shouldn't include babies with illness. They should be term if what you want to know is, is your formula going to be adequate for term infants. You don't want to increase the noise in your experiment by adding preterm infants. And then the question of ethnicity, terribly important in international studies, may not be so important in the United States where you're comparing what seems to be similar groups, but that's a question that needs to be carefully examined for each study.

So the duration of study should be at least 84 days. I think that 8 to 112 days, which is why our data are mostly 8 to 112 days, or 14 to 112 days, almost as good I think. At 8 days, there are great advantages of 8 days. Most formula-fed babies have regained their birth weight. You can get a really good measurement of length at 8 days. You're not there to get it when the baby is born, and even if you are there, the hardest time to get an accurate measurement on a baby is at birth. Howard Meredith, many years ago, showed me some publications on how very hard it was to get an adequate length at birth.

Now, I think that in the current U.S. climate, where you can hardly recruit any formula-fed babies before 42 days of age, which is still fairly easy, but you may be able to recruit them at 28 days. Later than

28 days you miss too much of the postnatal growth spurt. Earlier it's too hard to recruit, so maybe 28 days is the most feasible, and 28 to 112 is 84 days, and that's how I came to 84 days.

Now, if I give you my recommendations to the FDA, straight out, no hedging, I would say that size data are not relevant. I mean I know half the people in the audience hate me. Size data are not relevant. Data over 6 months of age are not relevant. I'm not even sure that data over 4 months of age are relevant. Breast-fed babies are not relevant, but that's not so serious because during the period we talk about they gain about the same. I just don't like to muddy up a study of formula-fed infants with a mixed group that I can control. And unless you have--you must match the cohort with the reference group. If your reference group consists of term infants, then you can't muddy it up by including infants, preterm infants in any number that they might be present.

So those are my messages and I'm willing to take the flak because I've done it before.

DR. GARZA: Thank you very much, Dr. Fomon. I regret the miscommunication between FDA and the speakers. We'll try to get that resolved for you in terms of the 20 versus 15 minutes.

Are there any questions? Dr. Anderson?

DR. ANDERSON: Anderson. I understand that the Iowa Fels data are longitudinal data.

DR. FOMON: Absolutely.

DR. ANDERSON: And that the recently published CDC standard is largely from cross-sectional data.

DR. FOMON: Absolutely.

DR. ANDERSON: To what extent do the percentiles generated from the two sources differ in substantive ways?

DR. FOMON: I don't know that offhand. I think from the point of view of infant formula evaluation, that it's not relevant because you can't get good analysis of longitudinal data from a cohort under study by comparing with size data. In 1976, just to prove my seniority, I had a dialogue with Peter Hamill [ph] over about 9 months, trying to convince him to call the NCHS charts size charts so that people wouldn't be confused by thinking that they're growth charts, but I lost that argument.

DR. GARZA: Dr. Stallings?

DR. STALLINGS: Dr. Fomon, I'm interested in your opinion of a couple things. One, the need for a control group, a contemporary control group when you're actually doing a study, you know, particularly for the kind of things we would be looking at, a change in formulation.

And secondly, when you said that you thought the breast-fed baby didn't have any role in thinking about evaluation of formula-fed babies, I think rather than is breast feeding--the growth pattern of breast-fed babies a pattern that should be strived for with formula-fed babies?

DR. FOMON: Well, let's see, question number 1. Tell me question number 1 again.

DR. STALLINGS: About a contemporary control group.

DR. FOMON: I think there are circumstances under which a concurrent control is essential, and one was the study that I mentioned to you, where we had a formula, a low-protein formula fortified with methionine or not fortified with methionine. In that case we had to demonstrate that with methionine it maps the reference group. But I think in general it's probably not really essential if you have good reference data for comparison, don't think it's really essential to have a concurrent control. It greatly increase the number. The number that you need to compare one cohort with the reference data is considerably less than the number that you need if you're going to compare it with a concurrent control.

So just from the practicality of making it possible at all to study new formulas, I think it's a reasonable compromise not to require a concurrent control

unless something about the ingredient change or whatever suggests that a concurrent control would be valuable.

And on the other question, should a cohort fed a new infant formula be compared to a breast fed control reference group, I think that's more a philosophic matter than a scientific matter, and my own conclusion is that if you want to study a new formula, you should study it in comparison with old formulas, and not with some group that we think might represent ideal growth. That's just what I think. Other people think other ways.

DR. GARZA: Any other questions?

[No response.]

DR. GARZA: Thank you very much.

It is 5 minutes to 12:00. We're going to try to make up the 20 minutes from this afternoon's schedule because we've been running a bit over time in the presentations, so that rather than coming back at 1:35 as the schedule suggests, I'm going to ask people to come back at about 1:10. So that we can make sure we can started by that time, having everybody here at 1:00 o'clock would be ideal. So we will ask the Committee to reconvene at 1:00 and we'll get started after that as possible.

Lunch for the Committee and the speakers is in the room where you had coffee, and the cafeteria is available to everyone else. 1:00 o'clock.

[Whereupon, at 11:57 a.m., there was a luncheon recess.]

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

(1:07 p.m.)

DR. GARZA: We can get started. Our next speaker--can I have all our guests please take their seats? I think the Committee is at the table.

Our next speaker is Dr. Jon Tyson, who is a professor of pediatrics, obstetrics, internal medicine and epidemiology at the University of Texas Medical School in Houston and the School of Public Health. Dr. Tyson will not be able to stay with us tomorrow, so I want to make sure that the Committee members ask all questions or clarify any outstanding issues that they might have related to the topic of growth data for preterm infants, because in fact Jon will be leaving soon after his presentation. So it's important that you try to get your questions to him before he returns to Houston.

Jon, thank you very much for coming.

DR. TYSON: Thank you very much.

Well, as a neonatologist and epidemiologist, I'm going to try to make the case today that if what you mean by normal growth is desirable or healthy growth, that the evaluation of early growth in preterm infants will necessarily involve evaluation of health and development. I'm going to try to go through this in a orderly fashion to promote a rational and evidence-based decision making.

The first question is: can the growth rate sustained by a new formula be adequately assessed using published growth norms?

This is a growth curve from the Neonatal Research Network, observed rates of growth, are weight--sizes, Dr. Fomon would say--in babies according to their birth weight. Throughout their hospital stay there is serial information assessed. Also in length and head circumference and mid-arm circumference for these babies. There's no data for babies greater than 1,500 grams and no data beyond discharge. I would add that there in my little handout, there's a website that you can go to where you can, for an individual baby, enter the measurements at birth, and print out a growth curve for that particular baby.

For preterm babies beyond discharge, this shows what I think is the best available data from the Infant Health and Development Program which had 985 preterm babies, and it provides data for 3 groups: 2,000 to 2,500 grams, 1,250 to 2,000 grams, and less than 1,250, plotted according to post-conceptual age--I think that's really post-menstrual age--up to 3 years of age.

Now, do these both norms describe normal growth? And I think we have to be really careful what we mean when we say normal values, because I think that's a term that often causes confusion. I actually try to avoid

that word. I see it used sometimes as referring to values that are expected or typical, typical values, values that are not associated with adverse outcomes or low-risk values, values that do not cause adverse outcomes, healthy or optimal values, and values for which intervention has not been demonstrated to be beneficial. I think it's really important that we keep clear what we're talking about. At best these growth grids describe what are typical values.

How should they be used? I think that they're a useful clinical tool to assess the growth pattern of individual babies. Whether they are appropriate regulatory standard to evaluating the formula, I'm sure the answer to that is no. They provide neither the optimal values nor an adequate basis to compare the growth with new formula to conventional formulas.

As I've spelled out in the handout, if you compare the growth for the new formula to one of these growth grids, what you describe as an effect on growth might be due to any of a large number of factors, including intervening changes in care and outcome since the growth grid was developed, the fact that in most studies you're using selected patients, whereas the growth grids are based on all patients; a myriad of differences between centers, and the opportunity for bias in patient care, selection, care and assessment in

evaluating new formulas. And I think it's particularly important when the sponsor or the investigator has a financial or even a professional interest in the outcome of the studies to attempt to avoid bias.

I think we also have to ask ourselves whether the statistical tests that are commonly done in evaluating growth studies are misused and then misinterpreted in assessing interventions using historical controls. The babies in feeding studies and the babies in these norms that we see for preterm babies and for term babies, are clearly not a random or even a representative population of the sample of the same population. Moreover, a p-value of less than .05 is often taken to mean that the difference is due to the intervention, when it may be due to any of a number of factors a difference in population, differences in the way the populations are assessed, et cetera.

So I think we should be asking the question, why use historical controls at all to evaluate new formulas for preterm infants or for term infants as well. For every other intervention that we talk of in medical care, the randomized trial is the gold standard. The concurrent cohort, carefully done, carefully studies, is a silver standard. Historical control is a bronze standard. Why do we want to use the bronze standard? And these are also issues in concurrent controls as well.

Should carefully designed randomized trials be required? I think the answer to that is clearly yes with a number of other features designed to minimize random error or systematic error and increase the signal-to-noise ratio, which would include mass caregivers and evaluators, well-standardized evaluations shown to be reliable by the people who do the assessments in the study, effective procedures to avoid attrition, and intention to treat analysis predefine stopping rules in an adequate sample size. And I would add to that I think a commitment to publish the data at least on the website, no matter what the data show.

Now, there has been opposition to the use of clinical trials by formula companies that would be charged with this responsibility, and I think part of that opposition is the expense and the feasibility of such trials, and I think with the progress in organization of research effort, that this is more feasible now with lots of neonatal research networks out there, some that already include follow up evaluations in at least some if not all the centers, and by the recognition that you don't--you need only do simple management trials.

Most people, when they think about randomized trials are very expensive, are thinking about the usual traditional kind of explanatory trial, which are designed

to determine whether therapies work in ideal or restricted circumstances, or that are designed to define the mechanisms of action. Management trials, on the other hand, or so-called effectiveness trials, are designed to determine whether therapies work under routine clinical circumstances, so all the effort that goes into trying to control all the co-interventions in explanatory trials is inappropriate in a management trial.

Who should be enrolled? Who should be excluded? I think you want to enroll representative sample of the babies for whom the formula is intended or at least the highest risk group. So you want to include any babies who make up an important part of that population, small for gestational age babies, very sick babies, twins, et cetera. You would want to exclude relatively few infants, say the babies who have major congenital anomalies or overt nonbacterial infections.

Should infants fed their mothers' milk be included? I don't think it's absolutely necessary, but I think it's highly desirable to increase the generalizability of the results of the trial because a large proportion of all preterm babies are fed at least some of their mother's milk, and also to help identify limitations of the formula, and areas for potential advances based on the benefits of mother's milk over

formula after adjusting as best feasible for other factors.

What assessments should be performed? I think, obviously, body composition or biochemical, physiologic or functional variables need be considered, but most of the time those would have been studies in prior explanatory trials. There might be some need to get some of that in some of the patients though.

What about health outcomes? Well, I think one variable that has to be considered is the percent of infants with necrotizing enterocolitis. This is a serious disorder with a mortality exceeding 50 percent in surgically treated extremely low-birth weight babies, and it may well be related to feeding. Death is a competing variable for necrotizing enterocolitis. You have to live long enough to get NEC, so you would need also to look at the composite outcome of death or necrotizing enterocolitis. We also have to be worried when we feed babies with chronic lung disease that rapid growth may not be attainable or even desirable. If you have marginal pulmonary sufficiency, how fast do you want that baby to grow? So you would like to know about the combination of death or prolonged mechanical ventilation.

Neuro developmental outcomes I think are at least highly desirable if not mandatory, because first growth and development may be differentially affected.

You need to exclude adverse outcomes on development, even in the presence of good growth rates. I think also we need this information to better define the optimal growth rate and the appropriate goals for growth rates sustained by formulas for infants with or without serious illness.

While growth assessments, weight, length, head circumference and weight-length ratio I think are essential, there may be others as well. What minimum period of assessment is needed, I think we should remember that a reliable identification of major neural developmental impairments is probably not possible any earlier than 18 months adjusted age, that is, post term, and this would allow evaluation of potential late effects beneficial or hazardous on time-limited interventions given in the NICU or later.

What standard should be used in judging the growth of preterm infants fed new formula? The American Academy of Pediatrics has said that the goal should be to achieve rates in an extrauterine environment like those that would have been achieve in utero had the baby not been delivered early. But we have to wonder if this is really the right goal. We have some uncertainty about what this rate is currently and we can discuss why. If we can only measure gestational age well, we could do this better. Currently we think that it's about 15 to 17 grams per kilo per day weight gain, about 1.1 centimeters

per week in length, and about .7 centimeters increase in head circumference. We again need to ask is this an appropriate goal for infants with severe lung disease, and we also had this observation of persistent growth deficits after reaching full feedings. So even though we can get babies to grow rapidly, once they get to full feedings, there's this long period of time when they're growing poorly as they recover from illness after birth, and this shows you data for the neonatal research network for babies of different gestational ages, 24 to 25 weeks, 26 to 27 or 28 to 29 in relationship to a so-called growth grid that Alexander published. And you can see that the babies don't do that bad once they start growing and taking a full intake, but they end up with most of them smaller than the 10th percentile for babies developing in utero with the same gestational age.

Whatever goal we try to take we have to think of it as provisional, but we could ask, should the current standard for judging preterm formulas be the formula that sustains the best catch-up growth, and that could be that the weight, length and head circumference and the body proportions would be most like that of term infants of the same adjusted age, providing there were no adverse effects on the health or development through 18 months as identified in a well-designed trial.

How many infants would you need to study to assess a new formula? This is a really complex and important generic issue in assessing intervention, any intervention where there may be an uncommon but serious potential hazard like necrotizing enterocolitis, and I'm going to spend some time on this even though this may seem to you like a statistical issue only, I think it's an important practical issue, because the kind of things I've said would make feeding studies so large that formula companies or indeed the NIH may be unwilling to fund these. So I want to try and see if we can find some way to address this.

So somebody may say, well, wait a minute. The old formulas have not actually been tested that well, and I've got a new formula here that has strong a priori evidence and rationale for using it. Say it has a component that's provided before birth across the placenta and in human milk after birth, not given in prior formulas. It's not well synthesized from precursors in preterm babies. And we think it's important for healthy development.

I think even in that circumstance you still need to rule out the possibility that there are important unrecognized hazards of this formula and I'm going to try to list what I think those are or the most important.

The first would be an absolute increase of at least 3 to 7 percent or more in major adverse neonatal outcomes, particularly necrotizing enterocolitis. A 3 percent absolute increase corresponds to what's called the number needed to treat of 33. That is, for every 33 babies that you this formula to, you would cause one baby to have necrotizing enterocolitis. I think that would be unacceptable even if all the other babies benefited in growth or perhaps even in development.

A second would be a reduction in developmental quotient at 18 months of a quarter of the standard deviation or more. That's the mean developmental quotient of 18 months and reduction of a quarter of a standard deviation or more. If you observe that, that would substantially increase the number of preterm infants with a deficient or marginal IQ that would be eligible later for educational intervention programs. In the neonatal network this would correspond to a reduction of almost 5 points on either of the Bayley subscores.

Third would be a reduction of a quarter of a standard deviation in length or head circumference at 18 months. And this of course is arbitrary, but at least after recovery from serious illness, there's no apparent benefit of slow growth so that I would think a modest decrease in length or head circumference, if not weight, could be seen as presumptive evidence of harm. And this

would correspond in the neonatal network in 18 months to about 250 grams in weight, 1-1/4 centimeters in length and a half a centimeter in head circumferences.

Now, you may think that this is too small to look at, but I would reassure you that the sample size needed to assess necrotizing enterocolitis, if you use a sample size that's large enough for that, you can evaluate very small effects on growth.

Now, if you take a conventional approach to sample size, you would need 315 per group to have 80 percent power to identify a quarter of a standard deviation difference in either development or growth or size at 18 months, and an alpha error of .05, assuming you lose fewer than 20 percent of kids to follow up. The power to identify an increase in necrotizing enterocolitis would be 78 percent, for a large increase, 7 percent, that would be a doubling of the right of necrotizing enterocolitis in the neonatal research network. It would fall to only 22 percent for a 3 percent increase, so a really small power to look at a clinically important increase. If you said you'd like 90 percent power to identify a quarter of a standard deviation difference of 18 months, you'd need 421 per group. Your power to identify an increase in NEC would still be only 30 percent for a 3 percent increase in NEC.

What can you do about this? One potential way to address this is a non-inferiority trial. For the sake of time I'm not going to talk about that.

Another, and I think this is, practically speaking, the most attractive option, is to increase the p-value considered statistically significant in evaluating a serious hazard. As you know, the same p-value, usually p less than .05 is used for benefits and hazards in studies, and this is an arbitrary and not well justified practice. For a serious hazard like necrotizing enterocolitis, a higher p-value might be justified on multiple grounds. First is, we know in clinical studies the direction of bias is toward finding benefits rather than looking for harms. There's a lot more effort put into it in general, and the studies are powered to evaluate benefit rather than harm usually. But the hazard may be much more important than the benefit. And as pragmatic evidence, we know that data safety monitoring committees that review the accruing evidence in clinical trials will stop a clinical trial at a much higher value of p for hazard than for benefit.

The appropriate p value shouldn't depend in part on the cost of drawing the wrong conclusion. For a serious hazard like NEC, I would contend that we might select a p less than .30. That would still result in a 70 percent chance or higher that a difference of that

magnitude would not occur by chance under the null hypothesis. If you did this, what you're doing is you're increasing the risk of a false positive conclusion, that is, that you would conclude that the formula causes NEC when in fact it doesn't. In order to reduce the risk of a false negative conclusion, that is, a conclusion that the formula doesn't cause NEC when in fact it does.

So if we go back to the numbers we calculated before for benefit, at 315 infants per group, again, that was for 80 percent to look at a .25 SD difference, the power to identify an increase in NEC would be 96 percent very high power for a large increase, 58 percent for a 3 percent increase. So you're slightly more than 50 percent likely to identify it. If you use 421 per group, you then get down to a power of about 2 in 3 to identify a 3 percent increase.

If you found hazards at a p of .30 and benefits at a p less than .05, what would you do? Well, I think you wouldn't recommend the formula, you'd recommend further study, and that would be a departure from what has been done in usual practice.

I'm going to skip that one, and just conclude by saying that I hope I've convinced you that the growth of preterm infants should not be assessed in isolation from effects on health and development, that a large trial evaluating growth health and development to 18 months or

more is needed to assure that the benefits of any new formula outweigh any hazards in preterm babies and to better define the effects of different growth rates, and the growth rate that we should be looking for in deciding how to design preterm formulas.

Thank you.

DR. GARZA: Thank you very much. Questions or comments? Dr. Stallings?

DR. STALLINGS: It sounds like you've silenced us pretty well. You know, we at the beginning, talked about dividing up preterm from term infants completely, and it sounds like, from your presentation, that you really, in the area of growth, that there really isn't anything you learned from term studies that would influence you on preterm. Would that be a fair--

DR. TYSON: Well, I wouldn't say wouldn't influence you, but I don't think you can determine whether a new formula is appropriate for a preterm infant based on observations in term infants.

DR. STALLINGS: The other thing I'd like for you to elaborate on a little bit is I think it's often that-- you were telling us a bit about who the sample should be, and in essence, the inclusion/exclusion criteria. Which infants, if you elaborate, which infants should not be in a growth study? Which preterm infants should not be in a growth study of preterm?

DR. TYSON: Well, I think it would be babies with the kind of problem that's very unusual, and that's going to have an overwhelming effect on growth like trisomy 13, growth and mortality, nonbacterial infection. Other than those things, I think you're talking about 3 percent of babies or something. The rest of them I would vote to include.

DR. THUREEN: Thureen. Dr. Tyson, I know that in your paper you said you would include growth restricted infants as part of this because they're such a large portion of the population, but that you would substratify those infants for further analysis.

DR. TYSON: Yes, right. You can of course include explanatory evaluations within a management trial, so it might be that that formula has a different effect on those babies.

DR. THUREEN: In terms of neuro developmental outcome, do you think it would be fair to exclude infants who had had very high risk factors for significant neuro developmental outcomes, such as intracranial hemorrhage, prolonged asphyxia, evidence of white-matter disease, before the trial even started?

DR. TYSON: If the formula is going to be fed to babies with severe asphyxia, then I think you would want to test it in those babies. For some of those conditions you had mentioned, they would occur after you started the

feeding, so like cystic white matter disease you might not identify till 36 weeks post conceptual age or something, and that's really, that's potentially an outcome variable.

DR. THUREEN: Would you pair match those infants then at all with other infants who had similar risk factors or known disease that affects neuro developmental outcome, or would you just do a purely prospective randomized trial?

DR. TYSON: If you do a large randomized trial, first of all, it gets really cumbersome to try and pair anybody at birth. As long as you're stratifying by center and maybe a couple of other things like birth weight less than 750, 750 to 1,000, something like that, that you will end up with an approximately equal number of those babies in the two groups, and then you can go back and do an analysis. If you try to stratify for birth weight, SGA, gender, birth asphyxia, et cetera, you end up with so many huge strata that the study gets really complicated to do. And I don't feel as strongly about that as most statisticians, but my understanding of the school, most statisticians are towards the minimal prognostic stratification at randomization, and more toward post hoc looking at individual groups who should have been predefined ahead of time which group you were going to look at. Does that answer your question?

DR. THUREEN: That makes sense. And would you change any of your ideas about how to conduct a study if you are going to look at patients who this is their exclusive formula fed from initial feeding versus studies started when infants really attained full feeding? Do you think that it makes any difference on how you conduct the study if you're looking at those two issues? Did that make sense? Because yours sound like you're referring to infants who may start minimal enteral feedings with the study of formula, rather than waiting until they attain full feeding and then starting from that standpoint? Do you think it's preferable to do one or the other, or do all of your idea really apply to--

DR. TYSON: It depends on when the formula is going to bed. If it's a formula that's going to be fed in the first week or something like that, I think you want to test it as it's going to be used. Let's say that it was a formula that was recommended for us from the first feeding. Let's say for the sake of argument that it cause necrotizing enterocolitis, and you didn't enroll baby, and you didn't start that formula until 3 weeks of age or something, or at a point when the babies were on full feedings, you might miss that effect, so you want to test it as it's going to be used in the real world. Does that make sense?

DR. THUREEN: Yes. And then lastly, do you believe that there are really no good reference standards for growth in the preterm infant or at least a certain subgroup of preterm infants that may be extremely low birth weight?

DR. TYSON: In the Neonatal Research Network, we have research nurses that are doing standardized--that have done standardized evaluations of anthropometry in intervals. There are huge center differences. If you try to take the data from any center to apply it to another center, you could easily be mislead just by the center differences. So I don't see why you would want to use comparisons that would involve center differences or time differences. This was gathered data 3 years ago when they were using steroids, postnatal steroids more often or less often than they are now. Why not randomize and get the cleanest--I think the belief that you don't have to use controls, you don't have to use randomized controls, that you can answer the question with fewer patients is an illusion, that at a given number of patients your ability to get a unbiased answer to the question is going to be greater with randomized controls than with historical controls.

DR. THUREEN: Thank you.

DR. GARZA: Heubi?

DR. HEUBI: I think, Jon, this is all very interesting. I wanted to actually ask you a couple of questions, and you'd have to follow my line of thought here.

The number of subjects that you would entertain as being appropriate for a study is about 10 times what a typical current formula study would be.

DR. TYSON: Right.

DR. HEUBI: With that in mind, knowing what you knowledge is of the Neonatal Network, would the Neonatal Network sponsor studies like this because this is specific to preterm infants and it would be a potential rationale to study with partnering between industry and NIH money to do studies like this?

DR. TYSON: The Neonatal Network has a protocol review committee and standard procedures for--there's no reason that couldn't be proposed and seriously considered or accepted if it went through all those things. There are lots of networks out there. In Texas we've started a Texas network, and there's an Oxford network, and the Canadians have a network, and the Australians have a network, and I'm sure there's networks developing in the United Kingdom if they're not already in place. So there are a lot of people willing to do this. The hardest part is going to be the 18-month follow up. That's a lot harder than studying NEC, but as more and more people

say, "If I'm going to take care of babies this size, I need to know how they turn out. That means I have to have a really well functioning follow up system. So I think there are going to be people out there that can do it at much lower cost than if you just went to them on day one and said, "We're going to fund your whole follow-up effort in order to answer this one question.

DR. HEUBI: But I was looking at it from the standpoint of it being economically more attractive to industry to do studies through the network that exists through the NIH because some of the infrastructure already existed and was already being paid for in part by federal money.

DR. TYSON: Right. Kathleen Kennedy and I proposed to the network a feeding study, and one of the things--and I'm glad I was involved in this effort--the business about the p-values that I presented today, that was aware stumbling block to us when we got to the--the reviewers really liked it and the statisticians said, "Well, you're going to have to study 6,000 babies or something," and I was working under the same mindset I had been before, well, that's if you want to look at a p less than .05. But why not accept a higher p-value, or as I was going to say on the last slide, predefine what you think is an acceptable ratio of the number of babies who benefit to the number of babies that are going to be

harmed, and then test that and say, does the number of babies who are helped by improved growth or development, relative to the number who are harmed by NEC or worsening BPD, if indeed that's a hazard, is that an acceptable ratio or not? And try to design studies not to look at one outcome variable but at the relationship of one or two variables or perhaps even more. And I think that's a cutting edge issue in the experimental design that the time is ripe to do now.

DR. HEUBI: And this is a circumstance where clearly DSMB or some monitoring board, during the--

DR. TYSON: Yes.

DR. HEUBI: --would be pretty accepted.

DR. TYSON: Right.

DR. DENNE: Jon, you've made an argument about following preterm infants out to 18 months for neuro developmental outcome and I understand the rationale for that argument. How do you feel about similar studies in terminants?

DR. TYSON: Do you mean randomized trials or--

DR. DENNE: No. I mean the necessity for evaluating neuro developmental outcome at 18 months in studies of new term formulas?

DR. TYSON: I don't see why not. I mean everybody in this room thinks nutrition's really important. You take these formulas and you feed them to

millions of babies. Why test it in only 50 or 60 babies? Why allow it to go on the market without knowing does it have beneficial or harmful effects as far as you can tell with an appropriate sized study in term babies?

DR. ANDERSON: Anderson. In your discussion of NEC much of the calculations were done based on a baseline rate of 7 percent. How would you feel about some of the adverse event monitoring being done not necessarily in the context of a randomized clinical trial, but against some fixed standard? That is an infant formula would be unacceptable if it produced a rate of NEC above 10 percent.

DR. TYSON: There are several problems with that. One is there is a lot of institution variation in NEC and with any institutions, there are periods when the NEC rate really goes up, and when it comes down, nobody quite understands that.

And finally you have the potential problem of bias. There have been studies, for example, they took x-rays of kids with NEC and x-rays of kids thought not to have NEC, going to every pediatric radiologist in California, and they found this incredible variability in what was called NEC and what wasn't. And so when you're in the context of a study like that, I just think the opportunity for bias is there, and that your ability to

relate that to some other institution in a different point in time I just don't think is worth the effort.

DR. STALLINGS: Stallings. A slightly different angle. And we talked this morning a little bit about, or inferred a little bit about term babies growing too fast, and you know, that's of concern, but certainly for my clinical time, the worry over preterm babies growing too fast, and I can remember bedside debates about too fast and it's only fat and it's no brain and it's no muscle and all of that. I don't think we have nearly as good a handle on the body composition component of the preterm babies. And then you add to all of that the concerns that we're all reading more and more about, is early postnatal growth a very--I mean we know it's an important time, but a differently very important time in lifelong health. Would you make a few comments? I know you made the caveat about babies with really chronic lung disease and concern, and that's really related to CO_2 retention, but put those babies aside, and can we grow preterm babies too fast, and how do we determine when we're approaching that?

DR. TYSON: Well, I clearly don't have the answer to those questions. It seems to me the only way we can get it is to randomize babies to different feeding regimens that produce different growth rates and see who turns out to have the best health and development. That

might be a different answer for the really sick babies and the healthy babies.

DR. GARZA: In terms of randomized trials, the implicit assumption is that in fact it will be a concurring group of formula-fed infants. The other control group that's been discussed by this committee in using breast-fed infants, in this case perhaps infants fed human milk that may be fortified or enhanced with other nutrients, what's more important in your perspective, a randomized trial, which obviously you can't do unless you're willing to go down the hospital hallway and say, "Are you going to breast feed or are you going to bottle feed," which obviously is not feasible.

DR. TYSON: So among those whose mothers commit to breast feeding those babies to randomize them to the new formula versus the old formula for supplementation, when they need supplementation, so you can stratify pre hoc for intent to breast feed or not. Then randomize within those strata.

DR. GARZA: And then use the amount of formula you're feeding them as part of the co-variants in that?

DR. TYSON: Yes.

DR. GARZA: Any other questions?

[No response.]

DR. TYSON: Thank you.

DR. GARZA: All right. Thank you very much, Jon.

DR. GARZA: We're going to move from preterm infants to generic analytical issues. And Dr. Frongillo, who you've met before, is on next. I remind everybody Ed is from Cornell University Nutritional sciences.

[Pause]

DR. FRONGILLO: So I was asked to look at analytical issues related to the evaluation of formula, and I was asked to look at 5 issues, the sensitivity and usefulness of several types of comparisons, the potential for evaluating a meaningful difference in the growth increments per day, the impact of transformations of raw data measurements into normalized indices, advantages and disadvantages of comparing with various reference data sets, and finally, circumstances that might favor one type of comparison to another.

So the first issue is sensitivity and usefulness of several types of comparisons. In 1988 the Academy of Pediatrics had issued some guidelines for determining physical growth, for evaluating new infant formula, and just to remind you, the suggestion there was to look at weight gain over the first 4 months, to look at measurements at 14, 60 and 120 days, and to look at rates of weight gain in grams per day over the intervals that

are implied by that, 14 to 60, 60 to 120, and then the whole period 14 to 120.

Some possible comparison groups to infants who are receiving new infant formula would be infants who are randomize to receive a standard established infant formula or alternately infants whose growth is represented in a reference, infants whose measurements are in a currently available data set, and finally, another possibility, infants whose measurements are in a historical data set. So I'm going to go through each one in turn.

The advantages--I'll use pluses for advantage and negatives for disadvantages. So in terms of advantages for randomized clinical studies, is first of all theoretical close control of the factors that might influence the outcome, in this case, weight gain as an example, the incorporation of design feeders to minimize known sources of bias. This might be something like doing stratification based on certain characteristics of the infant, perhaps size at birth or something like that, and that the probability statement is justified--this gets to sort of statistical philosophy here--but that the probability statement is justified on the design itself as well as the statistical model that's used.

And these issues are important, especially the first one in light of the fact that we might expect that,

for example, if there was challenges in recruiting infants into a study, if it was difficult to recruit infants into a study whose caregivers are interested in formula feeding, we might be concerned about selection biases or the differential characteristics of people who agree to participate in such a study. And so being able to control for those would be important.

The disadvantage is that the sample size is going to be larger than in the other approaches. I've said here the sample size would be twice as large. That's under the assumption that the other data set that one might be compared, whether it's a reference or a historical control, contributes no sampling variability. It's even larger if we're willing to assume that, for example, a reference has no sampling variability. And then it might be as great as a difference of a factor of 4. So the sample size then in a randomized control study would be potentially much larger than if we didn't do that. If we compare it to a reference the advantages would be that we can compare to a known established reference, something we know and love. One group of infants would be measured. Therefore, saving on sample size and work. The negatives would be that the new cohort on new infant formula may be different in some important ways from the reference sample, and that the reference may exhibit somewhat different characteristics

in terms of the growth patterns from the new cohort. And I already showed you an example of that earlier when I was talking about WHO growth reference, and we know breast-fed infants grow differently than formula-fed infants, and we might be concerned about whether the reference is the appropriate sort of comparitor for the current cohort.

And in relation to the current U.S. reference, it's a cross-sectional reference, and if our interest is in growth increments, we would like something that would characterize the variability in growth increments.

And then finally, reference data are not free of sampling error. Reference data have a finite number of infants measured in any particular time, and if we're doing a comparison between a new cohort and a reference, we really should be taking into account the sampling error that's present for both.

And the other choices had to do with use of currently available or perhaps historical data. These are data that--existing data that are collected earlier in time, so I'll consider those together, obviously advantages if this minimizes data collection, but the disadvantage would be that the characteristics of this new cohort might differ from the current cohort in some important ways.

The second issue was the potential for evaluating a meaningful difference in growth increments per day. So I want to talk about what I call the smallest meaningful difference, abbreviated SMD. This is not the difference expected or what's previously been found. What it is, is the smallest difference that we think would be substantively important. It's a judgment about what do we think is the smallest difference that would matter. And this is inherently--we should be thinking of this as a population rather than individual characteristic. In other words, the population perspective is most salient. We're not trying to judge whether one infant is growing differently than another. We're trying to judge whether a whole group of infants, potentially a large group of infants fed on an infant formula would grow differently than a group or population of infants fed on some other means.

The previous recommendation from that 1988 guideline was 3 grams per day, which amounts to 318 grams over the 14 to 121-day period. And I tried in the background paper to sort of get some perspective on how big is this difference. So if we look at the increments in the Iowa and Fels data, this difference here is about the same difference in those reference data between the 25th and the 50th or the 50th and the 75th percentile, to give you an idea of how big that is. If we think about

high and low altitude, which is one of the factors that causes the largest differences in birth weight, differences at high altitude are in the order of 350 grams or so. So that difference is about the same or a little bit bigger than this.

And we're always concerned about the effect of smoking on birth weight, and this difference of 318 grams is about 50 percent larger than that difference.

Another perspective we can get is if we look at some prior results. These are some results from a paper by Roche, et al. in 1993. They compared for males and females growth on 3 different infant formulas. They also had breast-fed infants that's in the paper, but I left that out here for simplicity. And this is the growth that occurred from zero to 4 months, and the total number of infants was about 260. And so these are the values of growth that occurred during those periods for male and females. What I was particularly interested in was having us look at the differences.

So the differences among those 3 groups, for males were 210, 270, and 480, and for females were 110, minus 10 and then 100. So you can see that first of all the differences for the males are much larger than the difference for females, and these differences here are smaller than the 318 grams over roughly the same period, and this difference is quite a bit larger than that.

And so it seemed to me, just from looking at this, we might say, well, these differences maybe aren't too big, but my guess is most of us would be concerned about differences of around 200 grams over this period.

The previous recommendation for sample size that was needed per group for the smallest meaningful difference was based on 80 power, testing at .05, a one-tail test and this standard deviation. And the result that was given was for 3 grams per day, a sample size of 28 was needed. So I wanted to comment on this.

First of all, as Jon Tyson pointed out in the talk just before, we're trying to do what's like a bio-equivalence. We're trying to make a judgment whether things are really producing about the same effect or whether there's evidence that something is really different. In that regard we're particularly concerned about power, because if there was a real difference in growth, the power represents the probability that we would find that difference. And 80 percent power means one fifth of the time we wouldn't detect the difference that was really there. My guess is that's a risk that we're probably not willing top take. And so I'll argue that 90 percent power is probably better.

Also this used a one-tail test and there's already been comments this morning that we're probably concerned about differences in growth that are in either

direction, whether infants grow too slowly or whether infants are growing too quickly in terms of weight, so a two-tail test is probably more appropriate. And finally, this figure was based on a personal communication with one of the investigators at the time, and since then a larger standard deviation has been published. So if we go to 90 percent power, the same test for the p-value, a two-tail test and a slightly higher standard deviation, then I've given a table here. In the background papers there's even more variance of this.

But this shows that first of all at 3 grams per day, we would need 67, not 28 under these parameters. And then if we go, obviously, to smaller and smaller meaningful differences, then we're going to need larger and larger samples to be sure that we would detect them.

The third issue was the impact of transformations if we go from raw data to normalized indices. The idea of Z-scores is that we match measurements with reference values for age and sex. The Z-scorer then takes that measurement, compares it to the reference median for that age and sex, divides by the appropriate standard deviation. The primary purpose for Z-scores is descriptive. It allows combining together ages and sexes so that we can get one overall description of a where a population is or a group is. This assumes

that the pattern of growth in the sampled population is expected to be the same as in the referenced population.

In terms of the application of this to evaluating new infant formulas, first of all, age adjustment would usually not be needed if the measurements are taken at the prespecified ages, and I think Dr. Fomon gave us the motivation and the example we need as to how you do that in getting measurements within plus or minus 4 days, which is remarkable. If we can do that, then we really wouldn't need to do age adjustment. And then if age adjustment was needed--because we're all not that good--what we would probably prefer to do is include co-variates for age or interpolate and extrapolate the time series rather than converting to Z-scores as a means to deal with this.

And finally, as I've already said, males and females would typically be analyzed separately because there are sex differences in the growth response. So the need to do sex adjustment from Z-scores isn't even there.

The fourth issue was advantages and disadvantages of comparing with the various reference data sets. References are a tool to provide a common basis of comparison, and as I said before, the referenced population should reflect the growth that's expected for the children that we have under study. The reviews that were done by NCHS and CDC and also be the WHO in the

early '90s, led to the development of new references, and so we need to consider this in the context of those.

I've made a table here which shows what the possible references might be. The CDC 2000 reference has both breast-fed and formula-fed infants, but it's cross-sectional, which is a limitation in its application for this particular purpose. The WHO data has breast-fed infants. It's longitudinal, which is good, but besides having breast-fed infants, it won't be available for a while. The Iowa data and Iowa Fels data combine together breast-fed and formula-fed infants in a longitudinal, and so probably at the current time, if we had to do this today, we'd probably choose to use this as a reference if we wanted to make a comparison for descriptive purposes.

And then finally circumstances favoring one type of comparison to another, it seemed to me that we would probably want to consider using currently available data if we wanted to test several new infant formulas nearing us in time. If we were doing a set of tests over some period of time that's relatively close, then perhaps it would be efficient to sample from the same population for the whole series of studies, but without having to do repeated sampling for a comparison group. The concern always will be that the characteristics of later samples might differ from the earlier samples.

So in summary, this review suggests the for primary analyses, the main analyses to answer the question, that it probably is appropriate to choose a design with a randomized concurrent comparison group. Otherwise, we're not going to be really sure that the growth of the infants on the new formula would really be expected to be the same as from some previous comparison, and this is of a concern about the characteristics of the samples and the populations, and it's particularly concerned potentially because of selection issues as to who actually gets into and agrees to be in the study.

For descriptive purposes, it would be useful to compare the attained weight for all groups at each measured age with the current U.S. reference, the 2000 reference, and to compare for descriptive purposes, rates of weight gain with the Iowa or Iowa Fels data.

And then continuing the summary, the sample size per group of 28 is clearly, in my view, without sufficient power for meaningful differences of even 3 grams per day, and a larger sample size is needed even if we stay with that guideline. The smallest meaningful difference might be smaller than that. It might be perhaps 2 grams per day. I think some might argue for something smaller than that, something larger than that. That's something that would have to be considered, but in any event, this would imply a much larger sample size,

and determining the smallest meaningful difference should be based on the best understanding we have of the biology, but also on the required regulatory, clinical and public health decisions that are going to be made with this information.

And finally, to Anna Milman, who's an undergraduate at Cornell who helped me in the preparation of the paper.

DR. GARZA: Any questions or comments? Dr. Stallings?

DR. STALLINGS: Talk tome a little bit about the--we heard some of this in the other talk to, the differences in comparing just the needs and medians and in growth studies looking at the number of children that would be, for example, less than the fifth or less than the third, some cutoff that we also think of? I've been--you know, we tend to analyze the data just looking at the group means and differences, rather than a secondary analysis, or often we do these things that we don't have any idea because of the way the data presented, how many children, how many infants might have been in the tails of the extreme. Any thoughts on that from a statistical point of view or an approach to believing that a formula supporting growth properly?

DR. FRONGILLO: No, that's exactly right. This whole discussion that I've just given is based solely on

the idea that under a new formula as opposed to an existing formula, say, that what we would be looking for and all we would care about is shifts in the whole distribution that could be captured by differences in the mean. Now, if there are differences of other kinds like differences in the variability or that if there was concern that maybe with a new formula for most infants it does fine, but there is a small subset of infants that don't do well at all, this approach is not designed to detect that. And Jon Tyson's thoughts about that I think are important, in that if we're concerned about what's happened at the extremes, that's going to inevitably require much larger sample sizes.

We know that any time we, for example, categorize infants into whether they've done okay versus not done well at all, and we have binary data, the sample sizes go way, way up, because in essence we're looking for those rare, rare events.

DR. GARZA: Ed, a related question to the one just asked by Dr. Stallings. In a clinical study one might be interested in individual outcomes, so that one doesn't have to wait until the end to decide that the child did not grow well. What sort of analytical issues would this present when you want to be able to monitor for safety reasons the ongoing growth of infants, of individual infants?

DR. FRONGILLO: You mean during the conduct of a trial perhaps?

DR. GARZA: During the conduct of a trial.

DR. FRONGILLO: Well, surely, you would want to have some guidelines in place so that if a child is growing extremely poorly, that some intervention is made. If you're starting at 14 days, you're not going to want to wait until 120 days to intervene, to decide that this child just is not thriving at all. So that would have to be built in.

DR. GARZA: With that thought in mind then, you talked about perhaps 3 grams being--sample size of 28 for 3 grams being definitely inappropriate or 2 grams obviously more so. But that assumes a linear rate of growth through this period, which several speakers have described as highly nonlinear. Should we be looking at narrower periods and different amounts of growth during those narrower periods than 2 or 3 grams over that longer period, where one assumes then the linear pattern that doesn't exist?

DR. FRONGILLO: Yes, I would say so. And the 1988 guideline has suggested that, what was it, 14 to 42--now I can't remember--that intervening interval would be looked at. The standard deviations there are a little bit higher. The growth rates on average are higher, and there are also more variables. So if one was particularly

concerned about say the earlier part--if there were three measurements and one was concerned about the earliest period, he'd need slightly larger sample sizes based on that. But, yes, one would certainly want to be able to evaluate growth. And I think Dr. Fomon's admonition that it's the earliest--the time when probably most concern and maybe most sensitive to different formula characteristics may be the period when the fastest growth is occurring, and that would be the earlier part of that period.

DR. GARZA: And yet--with that idea in mind then, when do you have to worry about regression towards the mean? I mean, so that in fact if a child is born large, do they begin to downsize immediately, or is that not a phenomena that we should worry about in the first 3 months of life?

DR. FRONGILLO: No, I think that is a phenomena we should be worried about, and that ties to what I was saying about selection issues. If, for example, for whatever reason, the parents who agree to enroll their infant into a trial are parents whose infants tend to be small or have lower birth weights, for example. We would expect infants with lower birth weights to grow faster. That's what we know, and if they're at the high end of the birth weight distribution, they're going to grow more slowly, and so one of the important advantages of having

the concurrent randomized comparison group is if any regression that occurs because of that kind of selection, that will be common to both groups. Any other way of doing it won't be guaranteed to pick that up.

DR. GARZA: Any other questions? Dr. Stallings?

DR. STALLINGS: One last technical question while we have you here. Like Bert was saying, maybe the 2 grams across the whole period is an optimal. Should we consider different rates for the gender since that's a pretty well appreciated fact as well to do sample on boys versus girls, different sample sizes in designing the study, or is that not--

DR. FRONGILLO: I guess my initial reaction is figuring out what the smallest meaningful difference once is going to be difficult enough.

DR. STALLINGS: Forget it, okay.

[Laughter.]

DR. FRONGILLO: And now you want to do it twice.

DR. STALLINGS: I was just thinking about how if we were, if we got to design sort of perfect studies in the young group, the youngest age group, and thinking about what we heard this morning, where there's at least one good example where the boys didn't try it as well as the girls, and that if we had been in that unique place where it had been a balanced gender sample, we might have missed the boys. Anyway, I was just--

DR. FRONGILLO: That's right. I mean from data I've seen, from the Iowa Fels data the girls are slightly less--their growth rates are slightly less variable than the males. On the other hand, at least that example from the Roche, et al. data--others have far more knowledge than I do--suggest that males are more differentially responsive to differences in feeding mode, so if both of those are in place, that suggests that on the grounds of variability, a smaller sample size would be needed for females, and on the grounds of variability, a larger for males. On the other hand, the differences we see for males might be larger than the important differences we see for females. So some thinking about that is probably--certainly would be helpful.

DR. GARZA: Dr. Denne then Dr. Heubi.

DR. DENNE: I just wanted to clarify what you said about the growth of large infants and small infants, that larger infants grow more slowly and smaller infants grow more quickly. My understanding of the growth curves is that the 10th percentile and 90th percentile are fairly parallel in terms of their slope. So I guess I--

DR. FRONGILLO: On a population basis infants who are born at high birth weights tend to grow a little bit more slowly than infants who are a low birth weight, so you actually see a convergence of the distributions a bit. So that the variability at birth tends to be a

little bit larger cross-sectionally and little variability after that. It's not a factor that comes into play anywhere after the immediate neonatal period. But for this purpose that's relevant.

DR. GARZA: Heubi.

DR. HEUBI: You showed us an example of a circumstance of where the smallest meaningful difference was in fact probably too large. I guess the question I have is do we know from formulas that are currently on the market, how many of them would have fallen out if you'd applied more strict criteria to them in terms of their trials?

DR. FRONGILLO: I asked exactly that question at lunchtime, and at least the people at my table didn't know the answer to that.

[Laughter.]

DR. HEUBI: So we don't know the answer to that question.

DR. FRONGILLO: I think that it would be helpful to have an idea of how often is it that we see differences that are larger than whatever the smallest meaningful difference we happen to be fixated on at the moment, how often is it that when trials are done do we actually see differences that exceed that and therefore would cause worry? So knowing something about the distribution of the differences that are seen when

studies are done would probably be helpful. The answer I got was that's not something that has to be publicly reported, and so we don't have a database of information available.

DR. HEUBI: And I guess my follow up question was, I'm actually surprised that they agreed to a one-tail test to begin with when they actually made these recommendations. Do you have any comment about that?

DR. FRONGILLO: On why they agreed to that?

DR. HEUBI: Why they decided that one-tail T-test was adequate.

DR. FRONGILLO: Well, I think that if your concern was we have currently available formula that are fine and you're introducing a new formula and you're concerned about under nutrition, that it doesn't meet the infants' nutritional needs, then it would sense to have a one-tail test because you're concerned about deviations that are in one direction. I guess I'm suggesting that where we are now in the public health concerns we have about obesity, it's probably the case that we would be concerned in either direction. So at the time it may have been a perfectly reasonable conclusion that isn't necessarily the one we should reach now.

DR. HEUBI: I guess it's just been pounded in me so often not to use a one-tail test for anything, that it just seems like it makes no sense.

DR. GARZA: You don't have to answer that.

[Laughter.]

DR. GARZA: Any other questions?

[No response.]

DR. GARZA: All right. Thank you very much. You may not want to leave the podium just yet.

We have next on the agenda is an opportunity to bring all the previous speakers forward for a question and answer period. And what I'd like to ask the Committee is whether you would find that helpful or whether we should just move on to the next two talks and thus bring everybody up at one time. I need a sense from you as to what you would find most helpful.

On with the speakers? Is that the consensus. And then we'll just ask everybody to come up. Is there anyone objecting to that on the Committee?

[No response.]

DR. GARZA: If not, Dr. Benton? Dr. Duane Benton, who is retired. I don't know whether there is an emeritus ranking within industry, but certainly you would merit that title if it existed. So we will refer to him as an emeritus industry representative, at least as an industry employee in the past, but one of high distinction. And he will be talking about product composition considerations, clinical studies.

DR. BENTON: I want to make sure you can hear and other things. One, I'm not going to--at my age and my position, I'm not going to try to impress you with how fancy I can make slides. So we aren't going to have any. I hate to a slide, to a screen.

Also I think all I'm going the be doing is a critique of my own talk, so I've got some things here I don't want you to forget. I enjoy looking at the audience and if you're falling asleep, I'll start at you, something like that.

Now, you should have background of why am I up here. Well, one of my various qualifications is, as Bert said, I'm retired, and they somehow felt that would reduce the conflict of interest. Obviously one of the senior people. Sam's a little more senior than I am.

But in this, what I am drawing from is 25 years when I was director of nutritional research with the Ross Products Division of Abbott Laboratories. And when I refer to my company, then I'm referring to--some of you used to know it as Ross Labs back in the good old days, and it is now Ross Products Division.

It was a very good company to work for. It may not be representative of the industry in its attitude toward physical growth studies, and in fact what I say may not be representative of what the industry or the Ross Products Division presently do. In other words,

when I went into this, I asked the people at Ross. They were calling me up because Bert couldn't find me, because I was wandering around out in Kansas. And so they called me up and said, well, wouldn't you do this? And I said, "Do you want me to? I could be kind of a loose cannon."

[Laughter.]

DR. BENTON: And they said, "Oh, yeah." The people on the Infant Formula Council said, "That's a great idea." We'll see.

[Laughter.]

DR. BENTON: Because these are my opinions. This is not edited material. I really sort of identified when Benjamin Franklin was asked whether he would write the Declaration of Independence, and he said he made it an issue of point, that he did not write material to be edited by other people. I can understand that now that I'm in my--other than my wife did try to read it through and tried to figure out whether it made any sense at all.

Now, the issues in here are, in my 25 years at Ross I was doing all kinds of different things. I'm a biochemist by training. I worked in nutrition throughout my professional career, and I was very concerned about safety of our products. Now, in this we're not supposed to be talking about safety. We're talking about nutritional adequacy. When you are in the position that I was in for 25 years you don't know how to draw a line

between those. You remember back when you were a graduate student and there was SMA out on the market that was vitamin B-6 deficient? And that company was--well, it was a real blow. Rudy Tomarelli told me that he spent through his whole career going into court cases in relationship to that. That's the one thing you don't want to do if you are in the position I am.

So, safety, nutritional adequacy, et cetera, there is no discrimination in there. The formula has to be good, and you've got to somehow figure out whether it is or it isn't, and if it isn't, you'd better do something about it very fast.

I was in a company where the aspect of a physical growth study was something that was assumed. When we were in a meeting and they were talking with the president about, "Are we going to introduce this?" You know, the answer was, "Where are the growth studies?" You didn't want to be in the position of not having that kind of information because that was sort of viewed as something not right with it, and we're not going to get out there, and we're not going to be hit with something very undesirable.

So it's also very helpful to see what I'm saying here and what I've written in the paper, and I'm going to be talking basically to the paper, is directed toward that. Now, it is not an aspect of you saying, "Oh, well,

this didn't sort of turn out right and maybe we ought to do this and a little more of that." Within the company, clearly, they want to do something. You can tell them, "No way. This stuff doesn't look like it's safe." But tell them, "Oh, we haven't got that done yet, or "We haven't looked into that," or something like that. You better have looked into it. You better know what you can as far as the answers to it.

Now, what we're basically saying about a physical growth study, from our point of view, is that if there were a nutritional deficiency of any sort, some slight depletion of an essential nutrient, that the first thing we would expect to show up would be a diminution in weight gain. Now, the other thing that in many cases we would expect to see would be a diminution in intake of the formula. Dr. Fomon did address to a degree the intake of formula. Most of the other things are talking about growth of infants and things you can measure on infants and so forth.

In my evaluation, really, wherever I could, wanted to see if the kids really did take in the formula and would continue to do that, and if I could work that into a study I would. It's hard. It's very hard work, but there are ways to find out that you can do that.

Now, when we progress into--and you now, I'm talking about the studies as such. Before I talk about

the formula and the ingredients, although my job was formula ingredient, I've got to talk about the studies to the degree of what can we expect from them, because it isn't going to make any sense why I would want to run them if I could expect--if I didn't have in mind clearly what we can expect from them and what we cannot expect from them. And one point that I have tried to make in here, those things that we can measure in a physical growth study, and clearly you've outlined most of them that could be. I mean we could do innumerable measurements on these infants. But our attitude, or at least mine was then--I may find somebody that would disagree with me--are that growth is a combination of responses, thousands of them, and those might or might not be measured. But we need the most sensitive measure of how an infant is growing to detect in this short period of time, and with as small a number of infants as possible, that nothing is going wrong. And therefore we need the most sensitive, sensitive response within the study.

With that, one of the things that you people are addressing is, well, shouldn't we use a breast-fed group as the control? This just sort of knocks me off my seat to think of that. We know that those breast-fed infants grow differently than any other infants fed on infant formula, that there are real differences of what we can

measure as far as growth. And I'm trying to answer one question. I'm trying to answer the question, should we put this formula on the market? I may have made a bad infant formula, and now they do go over--they gain weight or length or whatever at the rate that a breast-fed infant does. That may be a bad situation. It's not because the growth of a breast-fed infant is bad. It is that I may have gotten there for a very, very different reason. I may have depleted a nutrient. I may have caused a metabolic imbalance. I may have done a whole bunch of things because you cannot equate a weight gain or a body weight or something like that in a infant and say, "These two are the same. Therefore, their metabolic patterns are the same. Their nutrient contents are the same." No, they're not.

If I get one that is like a breast-fed infant, I would probably conclude that there was something wrong with my formula, and that's basically what we do. I mean we look at all of the information that we can collect and is practical from that point of view. When you come down and you make a decision and you're going to make advice to your company, you're saying, "Is this safe when you compare it to a present formula that we believe is safe, and we have long experience?" If it isn't, you don't have a formula, and the issue is that you're going to go

back there and you're going to change the formula so you fit into that.

Now, clearly, there are innumerable research studies that can be done, and you people would probably look at the budgets that we've used for our research studies over the years because there are extensive ones, and you'd think, "Hey, that's a pretty good looking budget." That wouldn't even come close to trying to answer the question that you are posing in the sense of can we somehow make infant formula feeding similar to that of the breast-fed infant. That is an enormous research project. I'm not really sure how I would try to undertake it if I were forced to try that. I'm sort of glad I'm not.

The other thing I'm going to say here is that I'm not going to discuss the premature infant. Dr. Tyson gave a very good talk about the aspect of doing that. I guess I would say that if he were to implement and be able to implement his thing about the risk of necrotizing enterocolitis, all of our formula bottles would say the surgeon general has determined that this formula may increase the risk of necrotizing enterocolitis, because I have not seen a formula that in some way or other didn't appear in a particular nursery to increase the risk of NEC.

Now let's proceed on with this, and another thing that I wanted to address, and Dr. Fomon introduced me to that in his paper and in his talk. He said that these days it's going to be very, very difficult to start a study with formula fed infants before 28 days. This is very, very serious from my point of view. We really depended on Sam for our work, and he started those infants--well, it was not within the first week, but certainly the end of the second week. and if we had some kind of a nutritional inadequacy which represented the depletion of a nutrient, or I guess you could even say accumulation of a toxin. Those very early part of the growth would be a very serious loss of that period, would be very, very serious in trying to detect a deficiency, because sometimes around 4 months or 3 months or 4 months, solid foods are going to be introduced, the proportional growth rate is decreasing. If you had had 3 weeks of feeding of a formula or breast feeding, you might have depleted or repleted stores of a particular nutrient, and I'm not certain that I wouldn't miss a serious problem.

One of the things that I always fall back on is saying that when I was director I was referring to the vitamin B-6 deficiency in a formula. Would I have been able to detect that in a clinical study? I seriously doubt if an infant had been fed for 28 days on a totally

adequate infant formula that had been fortified with vitamin B-6, that at the 4-month period of time whether you would be able to detect a weight gain difference, you know, using that formula, because they would have received enough Vitamin B-6 that it pretty might likely have carried them through that deficiency, and we would have never known. And we well know what happened when it really went out on the market and infants were really fed on that product as the sole source of food from birth. So I think it will be a great loss if we cannot introduce our infants quickly into the formula feeding when we're testing.

Another thing that I would point out, at least in our studies--and it was tremendously important to our work with Dr. Fomon and most of our other centers--they had an expert nursing staff. They observed the infants closely, just as the pediatricians observed the infants closely. They talked with the mothers. They resolved observations that the mothers had and so forth. And I personally believe that nutritional or other types of problems will probably be detected by the nursing staff before we get statistically significant weight differences, which, from our point of view, would be important to be able to stop a study where something was going wrong. And we have had studies where--that were stopped, anyway. I don't know whether you would

statistically ever have found out whether something was going wrong or not. We didn't want to know.

One of the things that is in here and that I have not really discussed, but I would just want to mention in this, there are situations where claims are going to be made for infant formula. And, yes, if someone is making a claim for infant formula, a growth study is important. It's my personal opinion that there's an awful lot more research that is needed to go out and make some kind of a claim for an advantage or disadvantage of an infant formula. And no one has ever regulated that. I certainly would hate to see claims that I see at the health food store for other foods.

Now, when I went through here, I gave you on page 2--I tried to break down the aspect of, you know, how we make an infant formula, what the steps are in it, and then we'll try to discuss how those go into the aspect of our evaluation. Certainly the ingredients, which we have had discussions of, and they've asked me to try to discuss all of the other various ones.

But the aspects of things like batching, in other words, this is where you're going to put together all the ingredients, they have to be put together in a specific order, specific manner, specific temperatures, and specific times of what is going on in there, and it can make tremendous differences in the stability of

nutrients, interactions between nutrients, a lot of other things that are involved here. And we will say later on, you know--and someone mentioned before, hey, we could just test this ingredient in an infant formula someplace, and then everybody could use it. Well, I have serious doubts about that, and I will try to address those as I have time here.

Certainly heat processing is very, very important in the interaction between the various ingredients, and, therefore, we're going to have to always think about how much heat processing a particular product has received. But that heat processing is going to interact with all of the other aspects of things that I'm talking about here. And it has to do with, you know, aspects of how we get the product into the container, how it's going to hold up over shelf life, innumerable aspects of that.

Now, let's just go down through and look at a few things in relationship to proteins. Dr. Fomon addressed it. Maybe I better not spend too much time on it. But, clearly, if you think you're going to measure the protein quality of an infant formula as to its adequacy, you must feed a very, very low protein feeding. In fact, Dr. Fomon, just a little lower than you dared feed in those studies that you showed them here.

Certainly when you want to use a rat to study protein quality, you get half its maximum growth rate. I don't think you'd like that for a clinical study on infants. So if you want to answer about protein quality, you're going to have to do other types of research. There are such studies that are very, very complicated.

I also make reference in here to the PER. This is a rat study. From my point of view, it has almost nothing to do with how acceptable a new protein source is for the quality of an infant formula because the amino acid requirements of an infant, as far as, say, patterns of amino acids, however you want to address it, are very, very different, certainly must be very, very different between the rat and--but it is a very useful test. It is useful in the sense of if a formula is different, in other words, if I have done different heat processing, I have an ingredient that's interacting with the protein or something like that, it will tell me that something has changed in relationship to the protein in a PER study. And you better watch out. It is just like our attitude on the clinical growth study. It is a very sensitive test, but it is very non-specific. It doesn't tell you what is wrong. It says something is wrong. And that is very important to our work.

I gave you some interesting examples in here in relationship to why I think we're a clinical study, and

one of them I should mention because it got tossed into that, goat's milk. Goat's milk is a high quality protein. Why would somebody run a clinical study if you made an infant formula with it? Some of you even wondered.

Well, from my point of view, it is certainly different from bovine milk. It certainly can interact, as the protein mixtures always will--protein mixtures are complex mixtures, and they have a potential to interact with other ingredients, possibly to produce unsafe material or to change the quality of the protein. Our feeling would be that we would have to run a clinical study.

I tried to think through and look at examples of ones that you might look at in relationship to protein quality, and they posed the question of: What if you used ultra-filtration to clean something out of soy protein isolate? Well, I said that wouldn't really need a clinical study. It certainly doesn't damage the soy protein isolate. It doesn't remove anything that you normally think of as nutrients. It removes phytoestrogens and assorted other materials. And so I think you could justify that you didn't need to run a clinical study.

If we were faced with that situation, would we run a clinical study? Yes. I think so, Ross. You just wouldn't want to be sitting out there without it.

I also made some derogatory remarks about protein hydrolyzates. They scare me to death because it makes--and that's not the mild protein hydrolyzates. Some of you are smiling. These are the highly hydrolyzed mixtures that taste like mud and so forth. But they have a very definite use.

The interaction of those with other ingredients in the formula is unbelievable. Certainly you would run a growth study of anything, you know, that received that kind of treatment.

I think my time is running a little short. Let's go through--because you can ask in questions if you have anything about carbohydrate sources, minerals, et cetera. To me they are relatively straightforward vitamin mixtures.

Food additives is something that has bothered me, always has, the aspect of how you approve food additives. One of them--probably the only one that has been studied adequately for an infant formula from my point of view is carrageenan. I'm sure there are people that will argue with me about that. I spent my whole career working with carrageenan.

Other additives, however, are based on are they safe for adult population at a level that they would be fed there. The intake in an infant can be so much greater and where it's the sole source of--I put one example in here that I just wanted to show you of a product that is generally recognized as safe, or GRAS, as far as the Food and Drug Administration. It's considered to be perfectly safe to add to any food. FDA may disagree with me on that.

But when we went into using this food additive, we found out that we knew nothing about its metabolism. We didn't know that the infants were going to be excreting large amounts of organic acids that we knew absolutely nothing about. And so we started out with rat studies and with labeled compounds to try to find out what in the devil was happening to--in this case it's octenyl-succinic-anhydride-treated starch. And we found out that there were these organic acids that were being excreted in the urine, and I was proposing some rather extensive animal studies and hopefully to go into the studies later on. And after a while, the people in companies ended up spending quite a bit of money. They said, Do you think you're ever going to convince yourself these are safe? And they decided, oh, let's not keep dumping money into a hole.

So as far as I know, our company did not use that additive, at least in the way we were talking about. But it was certainly one that there was no legal bar to its use, and I felt that there was an awful lot that had to be studied about what the material was.

I also have the situation where at some time people came to me and said we think you have this horrible hydrolyzate formula, and we think it would be a wonderful idea if we could improve the flavor of it because the infants would consume more of it and grow more happily. Well, one, their food consumption was very good on it, even though it tasted--I mean, you would never drink it. But you probably would have a hard time when the formula bottle was open and be in the same room with it.

But, anyway, they wanted me to evaluate flavors and the possibility of feeding it to infants. And there are more than 2,000 flavoring compounds that are perfectly legal to use in foods, and I spent a lot of time looking over those. And where I could learn the information I needed about the metabolism and so forth, I had some serious reservations. Many of them I couldn't even find out what happened to the--what would--how they would be utilized.

I don't know that we have flavored any infant formula--I hope we don't, or I hope somebody around there

goes back and does that over again, because I wasn't too happy with it.

Now, I was asked to try to get criteria for the aspects of how you would evaluate whether a clinical study would meet--it's right at the end of the paper. There are five aspects that we have here, and I guess it at least shows that I'm a chemist because I feel I have to know the chemistry and the reactivity of anything that I put into an infant formula. And if somebody can't explain to me what it's going to interact with, what it's going to do, you know, I can't approve it going in there.

Then you start proceeding to how is the infant going to metabolize this and do we know anything about that. And I think we ought to have some data, even before you go into a clinical study, but have a reasonable idea of the metabolism.

Now, some of these things are easy. Sure, we use mono- and diglycerides as a multiplier in infant formula. Those are natural products of the digestion of fat. That's easy, both in the chemistry and the metabolism. But there are other things that are much more complex and sometimes require important research studies.

Then you've got to go through and evaluate your whole process, and this is something to get you to realize that the aspect of the clinical study is

something that's going to come last. All of these things are going to be done. All of the chemistry, all of the stuff that's done in the pilot plant, is going to have to be worked out before anybody thinks about, you know, feeding it to an infant. And I think with good reason.

But any change that we have made in the formula has to be thrown into the mixture and say how could it have changed some nutrient, some interaction of ingredient, or something like that, to damage it? We certainly have tremendously sophisticated chemists these days that can go in and tell us just an awful lot of things about what has happened when a nutrient is--the state it is in a particular formula.

And if we started to find out that it was very different from what was in existing formulas, I would have concerns and feel that you have to have it resolved even before you went into the clinical.

Also, you should know whether you have any historical experience with similar formulas. In other words, if somebody--you know, we need a new formula or marketing people have some idea that there's reason for one, and you adjust the amount of milk protein down a little and the whey protein up a little, and you, you know, shift a little in the amount of fat and some of these things, you can visualize, well, you know, gosh, we've got all these clinical studies, and we've done

things almost identical with that. Why do you want to run a clinical study? And it's right. There isn't any reason if you have historical experience in clinical studies or experience out in the market that a very similar formula was the same way.

But if you find that, no, it is very different, then you've got to proceed to these other issues and try to answer the question of, you know, do we have enough knowledge to not run a clinical study? It's asking a negative. And it has to be that way.

You also should ask yourself the question of whether there are going to be physiological effects, and as I note in here, people are always coming to us with marvelous ideas about some fancy carbohydrate that we can add or a new fat or something like that, and it's going to have some remarkable physiological effect. Well, it is?

From my point of view, that means that's it. We certainly must at least run a growth study. No way are we going to, you know, not have that. But we also ought to be able to within that study try to understand--try to measure something that will tell us whether it has that physiological effect, or any, and can it in any way be injurious? And don't tell me that, aw, but it's more like human milk. I do not think that human milk would be safe to feed to an infant if we processed it in the

manner that we process infant formula. I don't think any of you can make it into a product that would be safe.

Now, that's just something to remember in this. I'm not saying anything against human milk or breast feeding at all. I'm just saying we are tough on the stuff that we've put into those bottles. And, yes, I know you give some heat treatment to the human milk that mothers collect and bring into the hospital. But you don't give it the kind of heat treatment we give it. No way.

So these are the kinds of things that we have to do, and if you do get down through all of those and you're confident of all of these things, then you can draw the conclusion that, well, there isn't any real reason that you have to run a growth study. The likelihood of it is small. And often within our company, the consequence was, aw, run it anyway, even though it's expensive, because the last thing you want to do--as I say here, one should use caution in drawing this conclusion. Why? Because if you're out there with the product and it doesn't perform the way you want it to, life will be unbearable.

Can I answer questions?

DR. GARZA: Thank you very much. Yes, I'm sure you will.

[Laughter.]

DR. GARZA: Any questions? Don't make me--Dr. Heubi?

DR. HEUBI: Can you comment about pre-biotics and pro-biotics and what you would think a company would require in terms of proof that they're safe and potentially efficacious?

DR. BENTON: Whew.

DR. HEUBI: It would take an hour, huh?

DR. BENTON: Well, I've thought about the subject. It is scary. It probably could be done. You're dealing with types of materials that in most cases are going to be reactive, that you have no experience with intake of these materials at the kind of intakes that an infant is going to have. You really ought to be looking for various kinds--for innumerable different effects that you would have on the infant or could have on the infant.

I would have to have an individual example to try to deal with, and I'd probably spend the next month in the library, at least, because it--those are hard questions.

DR. GARZA: Dr. Baker?

DR. BAKER: You, I think, have made the best argument I've heard for not including a breast-feeding control in infant trials. But I still have a little bit of a problem with the issue of a baby who's growing too

fast. I think that we all agree that there are times when babies are growing too fast. How are you going to differentiate--how are you going to tell whether your formula is giving optimal growth as opposed to good growth?

DR. BENTON: Well, I think you can tell, one, I'm not going to. But even back in the days when, you know, I would have been involved in any decision within the company trying to address that issue, it would have been an awesome process because I really view growth as a very, very complex thing. I mean, it's not a maximum thing of how much fat do you have in the baby and, you know, what muscles. I mean, it's down to the cellular level. It is innumerable things. And to say what is the right growth, you're going to have to look at those things at least to some degree, and you're going to have to have follow-up for a long, long time, you know, to draw that conclusion.

DR. BAKER: I think some of us in this room would say that formula-fed babies are too fat or too big, and so we're trying to figure out is there a way to feed babies with formula just right. I think that's a real issue.

DR. BENTON: Oh, well, Dr. Fomon did a series of studies at different caloric intakes, and these are perfectly adequate formulas. They just have a little

more water added to them. And he was able to get down to a level that even scared him. He got down to 11 calories per ounce from 20 calories per ounce. And you've got a different growth rate. At that point he was a little scared, and we did quit on some of those things.

I was scared, too. But we could reduce the caloric concentration on the formula, and the way mothers would feed it, you would get a slower growth rate.

I'm fairly confident you probably would reduce the body fat some, although I'm not so sure of that. Certainly the relationship of growth to body fat and so forth is a very, very complex issue. Long ago, when I was at Cornell, we studied that in the rat, and we could do tremendous things, although usually the rat followed constant increase of fat increase with time no matter the crazy things that we did to the poor creature. That relationship seemed to stay constant. But you would have to address it and do that.

Of course, you could also reduce it by reducing the protein content of the feeding. I get scared of that, too. I guess you would say, hey, he gets scared awfully easily.

DR. BAKER: But you wouldn't claim that that was optimal growth?

DR. BENTON: I have no idea. I have no idea of what's optimal growth. I know that I have to sell an

infant formula--well, when I was out there, we had to sell an infant formula next week, and so we could not shut down and say, well, when we've got an optimum formula and an optimum feeding pattern and an optimum--got everything arranged, then we'll start, you know, introducing feeding again.

I think our company would be very--Ross Laboratories, my ex-company, would be very interested in participating in studies of that. The idea that as a stockholder in Abbott Laboratories I don't think they could afford it to fund the whole thing because it is an enormous project. But it might be possible to do it.

But if, you know, somebody comes to you and says, well, that's easy to do, I don't think they understand the situation.

DR. GARZA: We'll take one more question, if there is one more.

DR. DOWNER: Thanks for such an interesting presentation. Did I hear you correctly in saying that in your opinion you didn't think that clinical trials were indicated as this time for the matter at hand?

DR. BENTON: For what? Pardon?

DR. DOWNER: Did I understand you to say that clinical trials in your opinion would not be indicated for the matter that we're discussing at hand?

DR. BENTON: For formula we were discussing?

DR. DOWNER: Did I understand you to say that you did not agree that clinical trials were needed at this time for the matter that we're discussing? You didn't say that then? Okay.

DR. BENTON: No, I don't think so. Our comparisons would always be to an infant formula.

DR. GARZA: I'm sure we're going to come back to that during the general discussion, Duane, but let's move on to the next speaker. Thank you very much.

The last presentation before we move on to that discussion is Dr. Denny Bier, who directs the ARS/USDA Children's Nutrition Research Center at Baylor College of Medicine, where he is professor of pediatrics. And he will be discussing clinical consideration in determining the need for clinical studies. Dr. Bier?

DR. BIER: Okay. Duane said that the clinical studies come last. Well, here I am.

Whereas I'd like to make a few disclaimers, first, I was fed on evaporated milk formula.

[Laughter.]

DR. BIER: So if I say anything dumb, that's the reason.

DR. GARZA: Should we add that your Dad was six feet tall?

[Laughter.]

DR. BIER: Right. And my mother won the Nobel Prize.

Two, whereas I did nose around among my knowledgeable friends for opinions, although the ones I express here are my own, and Duane and Sam said they could say what they want because they're retired, and I'm a director and my faculty thinks I'm retired, so I can say anything I want here. Let's see. Forward, if I can figure it out. The round button? Okay.

General operational principles for this talk are that as far as I can tell, today's formulas all contain--contain all of the known essential nutrients, and we don't find nutritional deficiencies that we know about, and for this reason new additions to formulas are likely to be those which have other purposes besides simple replacement of nutrient deficiencies. And I have not considered things like--no, no, that's not it. Okay.

[Laughter.]

DR. BIER: I'm not going to press it again.

I have not considered these things which we heard about this morning. That was good because we weren't supposed to consider them. I did leave out one important thing here, which is political considerations, which are certainly going to enter into whatever happens here later, I would assume.

Okay. Real-life Investing 101. When the energy and metabolism bills are paid, Mother Nature makes a contribution to a growth fund. This is why measuring growth is so important, because it occurs when all the maintenance needs are met. So it's a fundamental indicator of adequacy. And many people have talked about this, but growth measurements are advantageous because they're simple, they can be measured accurately, precisely. They're not invasive. And, most importantly, they're non-specific, that is, they're the best generic indicator of the fact that something unanticipated may have gone wrong or may have happened.

Factors controlling linear growth are different from those responsible for body weight accretion. For this reason, I think it's important to measure linear growth. Now, we had some discussion of this in various ways this morning, but the genetics that control linear growth is different than the genetics that controls accretion of body fat. The hormones and things and growth factors that control linear growth are different. So I think if we're talking about growth as a whole, we have to measure both weight accretion and linear growth.

I do not believe that body composition measurements are mandatory for a clinical study today, for a formula study, because the relationships among body components and childhood--infant, childhood, and adult

outcomes are really not well established. And for that reason I would say it's nice to measure body composition. It's very important for research purposes, but it's not mandatory for routine clinical formula evaluation.

My next guideline is that animal studies are never a sufficient substitute for human growth studies. Several people have addressed this today. Animal studies are necessary for proof of principle, for preclinical assessments of safety, for all sorts of other things. But the rat stops there, or the pig or the mouse or whatever.

Species differences in all sorts of events that deal with growth and developmental characteristics, the hormones, growth factors, metabolic differences among animals, and the species-specific characteristics of milk components and how they respond to them just make it not realistic to substitute animal growth data for human growth data.

Guideline 4 is that the presence of a substance in human milk is not sufficient in itself to eliminate the need for a human growth study. First, as I opened with, current formulas are nutrient sufficient, at least as far as I'm concerned. The non-nutritive components of milk are not well characterized. Some enter milk passively. They come along with the movement of fluid and water and electrolytes. Many don't have any

established purpose, and some have profound biological activity but their role in human growth and development is not well understood, for example, all the growth factors in colostrum.

So another way of saying this would be that the non-nutritive components of breast milk, what their function is in humans today, many of those are hypotheses, and because they're hypotheses, they deserve to be tested. And then I'd just like to remind you that there are compounds in human milk that you don't want there. For example, the infant's largest source of dioxin is breast milk. Now, that's not the mother's fault. But the infant's least source of dioxin is infant formula. There's no dioxin in infant formula. There's dioxin in breast milk and, therefore, we shouldn't use-- I'm not trying to imply that anybody wants to put dioxin in infant formula. But it's just an indicator of the fact that its presence in human milk doesn't mean that it's necessarily safe or should be there.

Guideline 5, I don't believe that data from post-marketing experiences elsewhere are really sufficient to substitute for a pre-market growth study. Post-marketing surveillance is largely an uncontrolled, anecdotal experiment. The validity of post-marketing data is heavily dependent upon the reporting of adverse events, and I think there are well-established and

documented cases for reasons why people don't report adverse events. Sometimes they don't see them. They frequently attribute them to another causes. Sometimes they don't talk about them because of potential liability issues. And then they don't like to fill out the paperwork or get involved in answering questions about this. So there's a significant underreporting.

I also don't believe that post-marketing surveillance is likely to detect subtle growth effects because of the range of normal infant growth and development and length, some of the things we heard about earlier, and then the various confounding effects that occur in routine clinical practice, and these are not going to be helpful, I think, in understanding control differences in new formula.

Okay. Those are sort of guiding principles, and we were asked to provide criteria, and my first criterion is that an infant growth study is required if a substance is being added to human milk for the purpose of influencing human growth. Given all the things that are--just, for example, colostrum growth factors, these sort of things that might someday be added to infant formula.

If the stated purpose is because this material influences human growth, well, it seems to me you need a growth study. Now, this is a subset of a more generic question, and even though--I forget who this morning

said--these are not on the table, I'm going to throw in a few gratuitous comments. If one claims an effect or benefit for a material added to the formula, then you have to demonstrate the claimed effect or benefit. I think that should be a fundamental principle.

The Working Group on the Nutritional Assessment of Infant Formulas of the Committee on Medical Aspects of Food and Nutrition Policy in the U.K. recently issued a statement which said "the goal should be an hypothesized functional or clinical benefit based on defined outcome measures."

"Any modification which is hypothesized--this should be "or," I guess--"or claimed to have significant advantages should be subject to clinical trial." This is basically sort of the same generic principle.

Okay. The sort of complement to that principle, Criterion 2, is that an infant growth study is required for macronutrients or other components that are known to affect the hormones, growth factors, genes, or metabolites that control human growth. So we have, you know, a variety of examples of that, things that, you know, for example, growth--amino acids that are growth hormones, secretagogues. We have fatty acids that are known to have gene regulatory effects, things of this sort. And if that's the case, then one needs to do a growth study.

Criterion 3, I would suggest that an infant growth study is required for formula changes that result in nutrient levels that are outside of established ranges. Now, I didn't quite know how to say this and it make sense, so I'll explain to you what I was thinking.

First, I think it's time to update CFR 107 because it talks about, as I recollect, minimum levels of 29 nutrients, maximum levels of only 10. It exempts a variety of others, and I think we now have LSRO reports on term infant formula, preterm infant formula, and the new DRIs. And I think it's time to start establishing a new set of consensus ranges for term and preterm infant formulas. And that would also require addressing the exemptions that currently exist for preterm formula. Should it still be exempt? I mean, these are legal questions and regulatory issues. I don't know, but I think it's time to look at them again. And I think its time to address the issue of the various non-nutritive substances that people are considering or talking about may have benefit in infant formulas.

Once that's done, you would have consensus ranges. You know, they might turn out to be regulations. I don't know. But you would have some relative established ranges. If they became statute, then it would be illegal to actually manufacture or to sell a formula that was outside those ranges. But hopefully

science changes along the way or we advance in some way, and there will be a time when someone wants to test things that are outside of those ranges. And I think growth studies or studies of effect would be required for any proposed formulas that fall outside those ranges. And this is, in fact, you know, equivalent to the opinion of a recent workshop report from the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition.

Criterion 4, addition of an entirely new compound to formulas requires a clinical growth study. I also sort of slipped away from trying to define what I mean by an entirely new compound, but I mean things that haven't appeared in formula before.

Now, as I mentioned, substances that are present in human milk, in my opinion, is not a sufficient reason for an exemption from this. Substances not present in human milk, again, for reasons we talked about earlier, animal studies are insufficient for establishing either safety or efficacy. A human study is required. If a human study is required, it contains several endpoint variables. One is growth as the non-specific variable of overall adequacy or identification of untoward effects, and then some primary endpoint variables that are relevant to testing the hypothesis of why you added this material in the first place.

Criterion 5, all entirely new formulas require a growth study, and, again, I'm not talking about legal definitions here. These will have to become legal definitions, but formulas that are not simply modifications of products already marketed in the U.S. For example, if one has a marketed formula and one changes the composition in a modest way--and I don't want to try to define that here--that would not necessarily fall into this category. But if an entirely new company is formed that's marketing an entirely brand-new formula that's never been marketed here before, even though they've stolen it from Ross or Mead Johnson or somebody else and are now making it themselves, they would have to establish the proof of nutritional equivalency or--equivalency to or superiority to marketed formulas because of the things we just heard about from Duane, that is, the different types of ingredients, ingredient sources, production and processing variables, matrix interactions, differences in absorption, bioavailability, et cetera. This coincides with the 1988 Academy of Pediatrics Task Force position.

Then, finally, the last criterion is that formula alterations that are likely to affect GI function or nutrient bioavailability require a growth study, and it seems to me this is, again, sort of a fundamental axiom. I mean, the gastrointestinal tract is required to

assure bioavailability. And if you put something into the gastrointestinal tract that affects the function of the gastrointestinal tract, you have to show me that the function is still adequate. There are known matrices that are more or less difficult to extract nutrients from, changes in macronutrient composition that affect absorption, for example, different types of fatty acids, enterocyte function, nutrients that react with enterocyte receptors and, in particular, gastrointestinal flora that have effects both among what the flora do in the intestinal lumen and how they interact with enterocyte receptors.

Then, finally, something we didn't talk about at all here today that I recollect, things that might affect gastrointestinal motility that may affect the ability to absorb nutrients.

So, with that, I will close and answer any questions.

DR. GARZA: Thank you, Dr. Bier.

Any questions? Dr. Stallings?

DR. STALLINGS: Thank you. There were two things that I thought about while I was listening. Early on you were discussing body composition, and, you know, that's one of the things I think as a committee we're thinking about, the role that might play. And you made the distinction between body composition measures being

appropriate for research and not appropriate--and suggesting not appropriate or necessary for these infant feeding studies.

And I think as I've thought about this over the last, you know, few months that we've been doing this, part of it is what is the quality of the research that's going to be required to bring products to market.

So I'm not trying to put you on the spot. Clearly, you're hearing some of that, too, that what we might be doing in our laboratories under protocols that we derived and are those really--should those be different, we're trying to get at the same thing. So just maybe elaborate on that in that context.

DR. BIER: Well, first, if someone wants to add body composition measurements to studies for licensing of a formula, I think that's entirely appropriate because I think we need far more information about what body composition means.

Right now I don't know what to do with the data, so if you gave me an old formula, Formula A that's marketed and now you test Formula B and it has slightly more fat and whatever, I don't know what to do with that information. Does that make it better? Does that make it worse? I don't know. Certainly a lot of people would say, well, if the infant is slightly fatter, well, that makes it worse because they may be fatter later, and I

would say that's an hypothesis that we certainly need to test a little bit more. So I'm not sure I know what to do with the data. Once I understand from the research studies what to do with the data, then I think it might be required, you know, for a new formula. But right now I just don't the information's adequate.

DR. STALLINGS: But you're making a bit of a distinction between the research and the activities the companies are doing--

DR. BIER: Okay. What we call research--okay. You know, yes, I'm talking about research for a primary scientific purpose as opposed to research for a licensing.

DR. STALLINGS: Well, that leads nicely into the other question.

DR. BIER: Good. I'm glad I'm baiting you.

DR. STALLINGS: You did fine.

[Laughter.]

DR. STALLINGS: But the other thing you've heard us contemplate here is what is the role of the pattern of growth of modern-day contemporary breast-fed babies and looking at if not optimum but desirable, or whatever we want to call it, what are your thoughts on that, because we have had 20 years of--

DR. BIER: I don't know what optimal is, and I know if I want to test a formula and know that it's

adequate, I need to test it against other adequate formulas, not against breast milk, because I know it's going to be different than breast milk. And I think it was Duane who just said we may be able to produce formulas that are equivalent to breast milk, and until someone shows me that they are better in some long-term way, I don't know what that means.

DR. STALLINGS: Thank you.

DR. GARZA: Any other questions?

[No response.]

DR. GARZA: Taking Duane's suggestion that, in fact, because one can manipulate formulas in a variety of ways, theoretically at least, to be able to achieve a growth pattern that would be comparable to human milk, what formula would one use as a standard against which to judge, given the fact that formulas are evolving and changing over time?

DR. BIER: You mean any new formula with regard to another formula?

DR. GARZA: Yes. And if they're different, how do you decide if it's an improvement? Does following growth then become a moot point?

DR. BIER: If it's an established--if you're making a change in a formula that's compared to an established formula, then the established formula becomes one. If it's not, then I think what you have to resort

to minimally are growth data from, you know, formula that babies, not breast-fed babies--

DR. GARZA: Does that in essence then become the desirable growth pattern? And what evidence is there that that, in fact--

DR. BIER: I don't know what a desirable growth pattern is, Bert. Can you tell me what that is?

DR. GARZA: No, but--well, I was trying to follow up.

DR. BIER: Well, tell me. I don't know. I mean, I think that's what we struggle with.

DR. GARZA: That's right. The logic of the proposition that, in fact, if one were to use a formula, then you have to choose the formula that you're going to be using to compare it with. And if growth is going to be the criterion and that becomes the control or the standard, then, in fact, a priori, one has decided what is the control or what--a value judgment has been made.

DR. BIER: I think that's a fair question, and I'm not sure I, you know, know the answer. I think now we have a variety of infant formulas that at least by all nutritional criteria are adequate. I honestly don't know if we have the data sets that allow us to compare those to each other in these critical periods of time to determine what the noise is if we use any one versus the other. Perhaps the industry people can tell me that.

But that would be something that would have to be established.

DR. GARZA: The other implicit assumption is that one would use growth in isolation of any other information to be able to assess manipulations in a formula that would achieve a growth that's coming from human milk. One could use body composition or one could use growth rates at different periods, or one might be able to use metabolic indicators for adequacy to make sure that nothing was limiting growth in an adverse way.

Among those three, which would you think would be the most relevant or which others?

DR. BIER: Well, I was asked to talk about growth, so that's why I focused on that. I mentioned body composition because I don't think it's necessary. I think in the case of many of these ingredients, one would have to have a metabolic measurement. So if you're changing the iron content of milk, for example, you would certainly want to know what serum iron is or ferritin or something like that.

I mean, I think it depends on what you're adding, and you have to add other components that allow you to test what the functional changes in the formula are.

And, by the way, there maybe--you know, I think your question also has in it this generic issue of if we

re-establish what the minimum requirements are for testing formulas, which I understand the committee is now going to expand into those questions over time, yes, what besides growth is necessary and what are some of the fundamental things? We heard about neural development earlier. That's obviously terribly critical in formulas, and as you know, the latest additions to formulas, the LC PUFAs are in some degree based on changes that might occur in neural developmental outcome.

DR. GARZA: We heard about Einstein formulas last time.

Okay. Any other questions?

DR. KUZMINSKI: I listened to your presentation, Dr. Bier, and to Dr. Benton's, and I'd link--both of which I appreciated very much. And I try to link material that I've heard and read with the material presented in Table 1 that was presented to us right up front this morning and, coming out of that, conclude that there are very few instances where a change in formulation does not necessitate a growth study to validate the change effectiveness, the effectiveness of the change.

And I guess I reflect a little bit that this is going to cause--this is a gray area. I hear comments like depending upon the experience, internal experience of the manufacturer, the experience at the clinical

testing organization, that would conclude that a growth test is needed or not, is it a serious change, is it a significant change?

I guess I finally get to my question, that I suspect that this will cause difficulty for the agency to try and put a rope around to try and harness what is a significant change. What is a significant amount of internal experience that leads a manufacturer to judge that no trial, no growth trial is necessary?

I'd be very interested in your comments and Dr. Benton's comments on that observation.

DR. BIER: Well, you noticed I avoided all of that in my talk.

[Laughter.]

DR. BIER: No, I mean, that's the nitty-gritty. I mean, we're going to--there's going to be cases when you're doing something that's so outside of prior conventional practice that everyone will agree that you need a study. And then there are going to be minor modifications to formulas which I think most everyone in the room, including myself and people who think you need growth studies, will agree it's almost certainly unnecessary in this case if some information X, Y, and Z is established. That's where, you know, the FDA and the infant formula manufacturers and whoever address this will come to some functional rules. And I don't think

they're going to be very easy to find. I just don't know--you know, I don't think there's any simple rule. Well, the simplest rule is to say yes, any change of an infant formula requires a new study. That's the easy rule. But that's probably not going to be the functional.

DR. GARZA: All right. Thank you very much, Denny.

We're going to take a break now and come back at 3:45 and then have the speakers all come to the podium at that time for a more general discussion. Thank you.

[Recess.]

DR. GARZA: If we can get all the speakers to come up to the podium?

We will try to go no later than 4:30 with the panel--I'm sorry, no, that's not the whole committee. Sorry about that.

[Laughter.]

DR. GARZA: We are good, but I fear we're not that good. We now have everyone assembled, so let's start with questions.

Maybe we will done by 4:15.

[Laughter.]

DR. GARZA: Dr. Anderson?

DR. ANDERSON: Imagine that I've taken a marketed infant formula and added to it a new substance

and that the only information that I have subsequently beyond safety is that in a clinical growth study the children who were first measured at 14 days had a distribution which was set right at the median of the CDC 2000 standards, and when measured at 1 months, 2 months, 3 months, and 6 months the weight-for-age, length-for-age, weight-for-length and head circumference was all centered at the 50th percentile with--let's suppose that there were 100 such infants, and 2.5 percent of them were above the 97.5 percentile and 2.5 percent were below the 2.5 percentile at each of the measurement points. Tell me why that particular formulation should not be approved for marketing.

DR. FOMON: Because the ingredient that you've added was supposed to be added for some purpose, and if it didn't accomplish that purpose, there is no point in approving the formula. If you're going to add something to a formula, you have a reason for adding it. If that reason is not substantiated, I don't think we're going to let you do it.

Anybody else?

[No response.]

DR. GARZA: Do you want to follow up?

DR. ANDERSON: Yes. So a natural consequence of that would be that a double-blind, randomized, controlled clinical trial that wasn't focused on--that showed that

there was no difference in growth between the new formula and the standard formula, in the absence of any measurements demonstrating that the additive had a desired effect would mean that such a formula would be inappropriate for marketing?

DR. FOMON: As a member of the FDA decision team, I say you're right, it would not be appropriate to market it. There's no point in putting out just a whole series of formulas that are going to be marketed because of some proposed advantage if that advantage isn't present.

DR. GRUMMER-STRAWN: Let me address that in a couple of directions. First of all, you haven't given us enough information to say whether we would want to approve it or not, because all that we've looked at is the growth data. There may be other purported effects of this new additive that may have been demonstrated to be there. There may have been particular marketing reasons that the company feels is important. Those need to be evaluated, so there's a whole set of other characteristics that we might be interested in other than the growth data.

But with regard to the growth data, how do we interpret that movement along the 50th percentile as the group mean? Really getting back to the discussion that I had with regard to Bert's question, the question is:

What is the appropriate reference? What you have demonstrated to me is that this new formula leads children to grow in the way that they grow across the United States. Does that mean that it's a good formula? I don't know.

Suppose that the previous formulation prior to adding this new ingredient had children growing starting at the 50th percentile and then falling off to the 25th percentile, and this new additive now has caused them to have a more rapid growth? Is that a good thing? I don't know. Back to the question Dr. Baker was asking. As we're in a situation in the United States with a growing epidemic of obesity, starting in young children, is it a good thing to have faster growth in the first 6 months of life or in the first 9 months of life? I don't know. But certainly I would question whether the way children are growing right now in the U.S. is the optimal set of circumstances.

DR. GARZA: Dr. Bier? And please identify yourselves as you answer because whoever is recording this will have difficulties.

DR. FRONGILLO: Ed Frongillo. Just to add to that, I guess a concern would be that, for example, the 2000 CDC reference has an admixture of infants who were fed in different ways, and we know that breast-fed infants, at least who have been intensively fed for some

time exclusively on breast milk for some time, for a few months, grow quite differently. And the results that we looked at earlier showed that this has been replicated many times. In the Euro growth study, they show differences that are similar, not quite as large but very similar in pattern, and that's from a large, multi-country data set.

So I think one of the concerns--I mean, I think to me there's two parts to your question. One is you seem to be trying to push our thinking about are there any circumstances when an existing reference might be used as a comparator rather than, for example, a randomized clinical trial. And I think we should probably consider that part of the question.

The other part is then what's the appropriate reference if you're going to use a reference. I would say the 2000 CDC is probably not the appropriate reference because of the admixture of breast-fed and formula-fed infants. If there was a reference group of formula-fed infants, one could maybe take up your question and think about are there circumstances when that's a good enough comparator that you would avoid the trouble of doing a randomized clinical trial. Well, I argued in my presentation that it would be hard to justify doing that because in this particular case, things kind of came out nicely. But when we do the

study, we don't necessarily know how it's going to come out.

DR. BIER: I'd just like to go back to our mystery formula.

You know, when you put a new ingredient in infant formula, you're doing a very big experiment on some number of children. And I think one of the fundamental principles of experimenting on children is there has to be a significant benefit-to-risk ratio or an insignificant risk-to-benefit ratio. Well, if you don't tell me the benefits, then the risk becomes infinite, or at least incalculable mathematically if you're dividing by zero.

You know, it's almost impossible to prove that what you're going to add is infinitely safe. So it has to have, you know, the potential for some risk, and until you give me a reason for saying that the benefit outweighs the risk, I'm reluctant to take the chance.

DR. BENTON: Duane Benton. I guess I should comment on it. I certainly agree, I think, with all of the things that have been said. There isn't anybody in the industry that wants to get an ingredient in there that isn't there for some purpose. I mean, that's just going out there and looking for trouble. Every ingredient is trouble, and you're certainly not going to

put something in that you can't find some kind of benefit for.

Now, if, you know, a regulatory agency were to look at this and think that some marketing people had cooked up something that was off the wall for a so-called benefit, I think they'd be highly justified in saying, Forget it, this doesn't even justify a clinical study. Don't come back to us and talk to us about it until you can tell me why you're adding it and what the benefit is.

DR. GARZA: Dr. Stallings?

DR. STALLINGS: To come back to one of the issues I think I've been trying to get us to think of, and as our consultants, I want to pose it in a little bit different way. I'm still intrigued by the idea of what the standard should be and knowing that none of us at this point in time know what's optimal. But if we weren't in the current environment where we didn't have the history of the current regulation and the history of a very successful program of, you know, the additive tests, the previous formula, and it was 2002 and we knew the epidemiology of growth in this country, if we were just starting from scratch, what would we think of as a good way to say what the normal pattern of growth is? Because I think we're struggling with two different things. One, we have a history of something that's been working pretty well, and in many ways it protects the

population, and it's given industry a framework. But part of what we're being asked to do is to come back and really potentially look at it very freshly. And I think that's where I keep coming back to is the pattern of healthy children born to healthy mothers who are exclusively breast-fed, say, for the first 4 months of life, is that something as experts in pediatric nutrition and growth and statistics and--you know, is that--would that be a place to start if we didn't have the history that we're currently working on, and the pros and cons of that. I think it would be helpful for the committee to hear some of this.

DR. CHUMLEA: Okay. I'll take a whack at that. You said--

DR. GARZA: Please identify yourself, Cameron.

DR. CHUMLEA: I'm sorry. Cameron Chumlea. You said, okay, basically in 2002 we'll start from scratch or we'll reinvent the wheel here. Okay. And then from whatever information we collect, we would be able to determine what optimal growth is. Is that--

DR. STALLINGS: I have a standard--

DR. CHUMLEA: I have a standard which we could do. Okay. We could do that, and we would find out basically pretty much that's what we did with NHANES III. It was a study just recently done of children, collect information, gives us status information.

And if that was where we started, we wouldn't know that we have a problem of obesity because it's the previous studies that kind of gave us that. So we'd be starting off with potentially false information or bad information.

DR. STALLINGS: Four months. First 4 months of life.

DR. CHUMLEA: Okay. In the first 4 months of life. So if we want to do this again, it will give us what is currently right now. Okay? If we did this in 1940, we would have gotten what's then, or 1929, or whatever. Each of those would be what was at that present time, and they're correct for the present time. But whether that's really the best or the end to meet all ends, none of us particularly know that.

So if I was going to want to do that and to come up with the information and assuming that the federal budget is not an issue and we have plenty of money, you could do this by creating a sufficiently representative sample of children, multi-ethnic, both sexes. It would be very intense to collect the information. And we could collect, starting from birth and a multi-center study, the information that would give us both status and rate of growth from birth to 6 months. But you're talking lots and lots of money--well, relatively.

DR. FOMON: Okay. It's 2002, and we're throwing out everything we thought we had and starting over. And now we have ease of recruiting breast-fed babies. We know that they don't grow appreciably differently than formula-fed babies, providing we start at 14 days, because it takes the breast-fed baby a little longer to get on that post-natal growth spurt. So if you start at 8 days, then the breast-fed babies grow faster between 6 and 14 days because they're still in catch-up.

So we start at 14 days, and we compare breast-fed babies and formula-fed babies, and there's nothing too much the matter with that. And if we demonstrate that we can--that the growth of breast-fed babies now is the same as it was during the past 20 or 30 years, we can use our big cohort of reference breast-fed babies. We have to--if we're going to do it for a 4-month study, then the breast-fed babies presumably should have some rules like they're all solely breast-fed, and the formula-fed babies should have some rules like they don't get anything else except formula. But that would be feasible. I mean, I can't argue against that just because we're comparing formula-fed with breast-fed babies. For 4 months--between 14 and 112 days it really doesn't make that much difference.

DR. GRUMMER-STRAWN: Larry Grummer-Strawn. I think that the appropriate reference is largely as you

described it, that I think we need to have a real prescriptive reference. I disagree with Cameron in terms of we could do this based on going out and sampling the U.S. population, not only because of the problem of obesity but because we have an admixture that Ed was talking about of a variety of ways of feeding, and we cannot conclude that somehow mixing all of that together is going to give us what's optimal.

I think that in developing formulas, if we're starting from scratch and we don't have them in front of us, if we're developing formulas, we would want to develop those formulas to best match what Mother Nature offers us and what we have defined at this point in time as optimal feeding.

It is true that 50 years from now we might think differently of what is truly optimal and we will have to re-evaluate that decision. But that is the nature of public health, that we always work with the best information that we have available to us. So I think the appropriate reference is to find the population that is feeding in the way that we currently think is the best way to feed.

How we get there from where we are now I think is a very difficult question. I think we've raised interesting issues in terms of we already have a formula that is producing a different kind of growth. It is not

automatically better to--any formula that moves in the direction of breast feeding is the right formula. So we can't immediately assume that that is a better formulation. But I think that somehow we need to figure out how to get back onto the track that Mother Nature would have us on.

DR. BIER: Dennis Bier. I thought Sam was going to give us the history of this in the 20th century and it would save my guessing at it, because he was there, you know.

[Laughter.]

DR. BIER: I only guess. And my guess is we started out the 20th century with infants who were formula-fed, you know, not doing as well as breast-fed infants. And people developed the formulas, and, in fact, what they used as their standard was the health and growth of the breast-fed infant. And at some point in the middle of the century or slightly thereafter, they found out, yes, by God, we can make infants grow just as well and be just as healthy as breast-fed infants, in fact, even a little bit better.

Now, if we weren't concerned about the hypothesis that these infants have more fat, maybe, and are getting fatter on that basis later in life, the hypothesis that this is true, what would be arguing about here as far as breast and formula feeding? They're as

healthy. They grow slightly better, if not the same or slightly better. So, you know, our concern here about too much growth is based on accretion of body fat. And any standard relative to that is going to be an arbitrary number based on someone's belief at this moment, in my estimation.

DR. FRONGILLO: Ed Frongillo. Virginia, if I understand the question you were asking, if we could start over and dismiss the 20th century and we're now in the 21st century, what would be an appropriate reference? But then I think that begs the question of: For what question, for what purpose? What is one trying to evaluate?

For example, if our only concern is are the nutrients adequate and we're not really interested in growth itself but we're interested in the fact that growth represents the summation of a whole series of processes that include having adequate micro- and macronutrients, that's one question. That's a very different question than does it lead to growth per se that we think is optimal. And even that could have a time frame of the first 4 months, say, or first 6 months, or it could have a much longer time frame. I mean, there is some evidence--I don't think it's terribly compelling yet in its magnitude, but there's certainly evidence that

exclusively breast-fed infants tend to be less obese later one.

Now, one could pose a question that we would like to have the growth that sort of represents not just satisfying the nutrient needs of the infant and some pattern of growth that we think represents optimal for now, but something that tells us something about what's going to happen with this infant in the future. And so how we think about will probably determine what choice we make of the reference.

I think Sam is right in the sense that there are certain windows--if you took perhaps the mid maybe 1 month to 4 months--I don't know exactly, somewhere in there--you'd probably find on average that the amount of growth of breast-fed infants and formula-fed infants, the growth that's accrued during that time is about the same. They probably got there differently over that period. If you extend it to 6 months, you'd probably find out they're not quite the same.

One of the things that we should recognize is that there's been lots of studies done now that have shown that breast-fed and formula-fed infants don't--their pattern of growth isn't the same through the first year. Everything we know tells us that by 24 months they've come back together. But those studies have been done. A lot of that has been motivated by--and I should

have said this when I talked about the WHO Multi-center Growth Reference Study this morning--has been motivated by the management problems that have resulted from this discrepancy and pattern so that the concern has been that infants who are breast-fed appear to be faltering relative to the current international previous U.S. reference. That management problem has been a major motivator in going ahead with a breast-fed based reference.

So many of the studies have examined the differences in patterns in light of that question. Here we're asking a different question. And it may be that some effort to evaluate perhaps the Iowa data that have infants from both feeding modes might tell us something about this question that we've kind of overlooked in asking a different question.

DR. GARZA: A related question that is useful for the panel to discuss relates to Virginia's question. Given the fact that we don't have long-term data to try to assess the functional consequences of different growth patterns in the first year of life, where should the burden of proof lie in terms of should we assume than that differences, until proven otherwise, are significant? Or should we assume that differences, until proven otherwise, are not?

DR. FRONGILLO: Can I just ask, for my benefit, at least, for clarification? Differences from what to what?

DR. GARZA: Growth patterns, the first year of life, breast versus formula, I mean, or one formula versus another formula.

DR. STALLINGS: But you're getting into long-term--

DR. GARZA: Long-term--given the fact that we keep referring to the fact that we can't--we don't have long-term data that try to assess the functional consequences of growth differences in the first year of life--

DR. CHUMLEA: What's long term?

DR. GARZA: Seven years, 15 years, 30 years, I mean, however you choose to define it, there are data, for example, now suggesting that certain growth patterns may predispose children to obesity.

DR. FRONGILLO: But I think--I'll turn this to somebody else who has an answer, because I don't, but--

[Laughter.]

DR. GARZA: We agree to do the difficult, the impossible, but we're contracting out to you guys.

[Laughter.]

DR. FRONGILLO: But it seems to me that question you're asking, Bert, there's at least two aspects to the

scope of that. One could be--has to do with formulas versus breast feeding, and if we started again at the 21st century, would we ever agree to have infants fed on any formula? That's a very different question than saying given that we've already accepted that infants can be, should be in many cases fed on formula, about differences among the formulas. Those are different issues.

DR. BENTON: I may be sort of throwing a question back at you instead of trying to answer it. But I am really perplexed in the sense that you people seem to think that what you are seeing in difference in growth patterns has something to do with nutrition. And I have no confidence in that at all. I mean, clearly, breast feeding is an interaction between the mother and the infant. And what the infant consumes is very determined in that process. And if that is the case, what is changed is in the breast-feeding process and how it changes the baby and how it changes the mother and her outlook on feeding and so forth. Our fooling around with formula to try to duplicate that is utterly hopeless.

Now, clearly, if you want to have studies to do this, you're going to have to have randomly assigned feedings. And you know--

DR. ANDERSON: Yes.

[Laughter.]

DR. GARZA: Are you saying, Duane, just so I understand, that you think it's due to feeding behaviors between--differences between the two groups?

DR. BENTON: Because [inaudible - off microphone] I mean, they're not at that critical [inaudible]. It is later you're seeing the striking differences. And that is the point at which, you know, it's the mother and the infant and how they're interacting. I guess the other way you study it is you trade in [inaudible] mothers.

[Laughter.]

DR. GARZA: You think that would be true cross-culturally, I mean, the fact that these differences are seen in all settings, that, in fact, it still reflects behavior. Okay.

DR. FOMON: Fomon. First time. Bert, in answer to your question, and putting it, if I may, in the context of testing a new formula, if we don't know the consequences of more or less rapid growth during the first 4 months, all we have left is the comparison of how they grow in the first 4 months. And that's all we can do now. Either they grow as we anticipate on the basis of a cohort of breast-fed babies or a cohort of formula-fed babies, or they grow the same, and that's all we can do. There is very little in the literature that tells

you how growth during the first 4 months relates to any subsequent time.

DR. GARZA: Let me follow up because you've made an important point. Let's assume, then, that one can--given the fact that in Duane's presentation, he clearly pointed out that one can manipulate growth in a variety of ways by formula changes. How useful would body composition be in assessing both differences as well as similarities in growth pattern given what we heard in Dr. Ellis' presentation and DXA? Would that be of any use in helping assess the nature of the differences or similarities, or do you feel that that would be either trivial or not very useful?

DR. FOMON: I think it would give you very little information at considerable cost in what you would get from measuring length and weight, if done properly.

DR. FRONGILLO: Ed Frongillo. Comments to add to what Duane said. I think that we know--Bert knows this more than just about anybody--that human milk is a biologically active substance. There are regulatory factors there that we don't really understand very well. So in addition to the perhaps maybe behavioral differences, it could be that some of the hormonal substances in milk play some role in determining the pattern that ensues after that.

And then in response to your last question, I agree with Sam that for the kinds of questions that are being asked, it seems that having both weight and length information would be important. Whether--

DR. GARZA: No one argues with that. Body composition--

DR. FRONGILLO: And whether or not there'd be any additional value of body composition on top of that, I would be skeptical that that would bring much more information relative to the work that would be required.

DR. CHUMLEA: I'd just add, going back one question, to the effects within the first 4 months on some subsequent event, since we have recorded information rather than anecdotal information, and the issue, say, 15 years ago was that babies became fat adults type thing. In the analysis of Fels data, there was no indication that anything before, in terms of skinfolds and weight before 2 years of age had any impact upon subsequent levels of obesity by 5 or 6 that did affect, and, of course, Sun's been happily able to show that changes in weight stature, BMI stature are very predictive of--in childhood and adolescence are very predictive of subsequent risk for obesity in adulthood, but nothing that shows up within the first 2 years. Going to kind of what's called the Barker hypothesis, we've got blood pressure. We've looked at birth weights. There's

nothing within Fels where we actually have recorded information and any information of adulthood that has any link whatsoever.

From a nutritional standpoint, we actually have seven-day dietary records that were collected from the children. What we have noticed in the preliminary analysis of that is that nutrition potentially in late childhood, say 8 to 12 years of age, and early adolescence, 12 to 16, does have predictive values on bone mineral density in terms of milk consumption in adulthood. But nothing before that has any relationship--and we've got dietary records back as early as 6 months of age. So nothing there links up. So I think frankly that the infant is extremely plastic and can absorb a lot of insults or whatever that goes on and nature has it so that pretty much it can take whatever abuse we're going to give them, and they're still going to probably turn out pretty good.

[Laughter.]

DR. ELLIS: It's just like we're saying, that we don't know what the translation is at this point for babies between, say, in the first 4 months or 6 months of life really what that translation is down the road. We do know that, like Cameron has said, when you get to age 3, that there are relationships. Of course, the closer you get to adulthood, the stronger the correlations would

be, which is understandable. Of course, again the 3 to 5, whatever, the main factors going between the first 6 months and there, environmental, genetic issues, even behavior issues and so forth, that are going to compound a simple relationship between what goes on the first 6 months and later.

However, most people do argue that unless there's something unusual about a child, they tend to track along the same percentile. And so presumably then if you have a high percentile, in terms of being percentiles, you're a high fat baby or a high probability of any increased fat, and, therefore, presumably those are the ones who also appear in the higher percentiles later on.

Again, today can you say for certain a baby was 30 percent fat versus one who's 15 percent fat, that he or she has a higher probability of developing cardiovascular disease or obesity or diabetes? The answer is no, we don't know that.

DR. SIGMAN-GRANT: Sigman-Grant. My two questions are going to just demonstrate my ignorance, but be that as it may, nobody mentioned head circumference. Is there any difference between, say, formula-fed and breast-fed babies in relationship to head circumference?

DR. FOMON: Head circumference in a normal baby is proportional to length. And head circumference--

increase in head circumference is proportional to increase in length. The value of head circumference is in detecting discrepancy between gains in head circumference and gains on length, which has neurologic, diagnostic implications.

DR. SIGMAN-GRANT: And my other question is: Is there any difference in organ weight betweens breast-fed babies and formula--and I know that there's only one way to look at that. However, I'm just curious because we talk about gross--you know, gross weight and gross percent fat versus--I know, but that's--the reason I ask that is because of the potential bioactive compounds in human milk and its relationship to stimulating GI growth and that kind of thing.

DR. FOMON: I can answer that question.

[Laughter.]

DR. FOMON: The answer is I don't know.

[Laughter.]

DR. FOMON: And I don't think anybody knows.

DR. GARZA: Although there are some scattered reports at least that have been done with sonography for thymus differences. But I don't know whether they've been replicated, and for thyroid, I think as well. But I don't think those have been replicated, and they've been done indirectly rather than by direct examination.

DR. SIGMAN-GRANT: Well, I bring that up not to be facetious. It's to ask a question. There's just another thing we don't...

DR. FOMON: Any other questions that none of us can answer?

[Laughter.]

DR. GARZA: Is their pay dependent upon whether they can answer these questions?

Any other questions from the group? This group will be here tomorrow, but we should not be shy. Dr. Anderson?

DR. ANDERSON: Supposing the development of the infant formulas containing very long chain fatty acids, no growth studies have been done, but there was evidence of the benefit of neurologic development at 24 months or 36 months and the additives were generally recognized as safe. Anybody willing to proceed without a growth study?

DR. BIER: No, because there is a priori evidence that fatty acids, particularly polyunsaturated fatty acids, can have effects on acosinoid (ph) metabolism. They have effects on activating fatty acid oxidation genes, you know, reducing fat synthesis genes. So we have a lot of a priori information that suggests we should do a growth study. So the answer would be no.

DR. GARZA: I certainly want to thank each of you for being so patient with this committee and its unreasonable demands of you.

[Laughter.]

DR. GARZA: But we will continue being unreasonable. I hope that each of you can join us tomorrow.

We're going to turn now--and you are certainly welcome to join us as we begin going through the seven questions. You all have these in your packets. I'd like to see if we could get through the first section on metrics for the evaluation of normal physical growth before we adjourn, and if possible, to take on at least part of No. 4 and 5.

Before getting into that discussion, though, there is one item of take-home work for each of us, unless you can figure out a better way of doing it. That is, if you look at Question 7, what I'd like each of you to try to do this evening--on your own, it doesn't have to be done as a group activity--is to--well, it can't be done as a group activity, anyway--is to look at some specific--pick one or two specific changes in infant formula that you think would reasonably be expected to be accompanied by a clinical study. That is Part B to that question. And then use that as a basis for trying to come up with general principles and criteria that led you

to that decision. Or you can do it in reverse, but if you can choose a specific example and guidelines and criteria, I think that if each of us brings those to the table tomorrow, then we might be able to agree on general principles or criteria that we would be able to answer Question 7 within the time frame that we've got.

In the absence of that preparation, I think it may be more difficult for us to get there.

You can pick an original one or go to Table 1. It's whatever--so that if I want to deal with it in less than the abstract--Roger?

DR. CLEMENS: Roger Clemens. To that point, Item 7, first of all, I was a 17-kilo baby at 12 months. And you'd probably say I should have been very big right now.

DR. GARZA: Well, you do look big.

[Laughter.]

DR. CLEMENS: I must admit, I am taller than anyone in my family, if that has any merit. Just to help with Item 7, the IFC, the International Formula Council, provided each member here on the panel with a diagram--a decision tree for documentation of nutritional adequacy. This was to have been mailed out to everyone, but I wasn't certain that everyone received a copy.

DR. GARZA: We received it about a week ago.

DR. CLEMENS: That's great, because I did not, Bert.

[Simultaneous conversation.]

DR. CLEMENS: --copy of that decision tree. Thank you very much. I appreciate the fine work that Jeanne has done on that point.

Also, just a comment, because Moski (ph) had made a comment from the CDC on examples that would not necessarily require clinical trials. Again, on the handout that was provided on the clinical protocol that has been followed by the IFC members for the last 20 years, fundamentally, you'll see that on page 6 there is a list of about eight items, of which under the current regulatory guidelines do not require additional clinical trials. That doesn't mean that the companies won't do them. It just says they're currently not required. I just want to bring that to light to everyone.

DR. GARZA: If some of you did not get the e-mail, you should let--

DR. CLEMENS: They were all distributed. It's just a case of e-mails--

DR. GARZA: If you haven't received that type of information, then either the e-mail system is not working--but it was sent out. I know that I received it about a week ago along with three or four other items on the attachments.

DR. CLEMENS: I appreciate that, Bert. I just bring out No. 7 in particular because you talked about the items, and I want to be sure I have that as a guideline because it talks about those changes, what changes would require clinical trials under this current environment versus those which would not.

DR. GARZA: I'm glad you raise that. I'm just more concerned about whether information committee members may not have received that was shared--

DR. HEUBI: I have this letter from IFC but didn't get this document.

DR. CLEMENS: See, that was my fear that they did not receive all the documentation. Thank you for the opportunity to share that.

DR. GARZA: All right. Then is dealing with Item 7 in the way that I've described acceptable to the group so I don't surprise you tomorrow morning? I'll turn to each of you and say no.

All right. Then with that housekeeping item, let's turn to Question 1. And we agreed that we would do this for term and preterms this morning as we reviewed the questions. It's asking us to group the following metrics in terms of clinical usefulness as endpoints, and I would suggest that we try a grouping scheme that says extremely useful or mandatory or whatever, of moderate use, C, of no use, or, D, it's still in the research area

and we really can't comment on whether it's of high, medium, or no use without additional research because there's enough biological plausibility or there's enough potential interest in the item that we would want to get it, but we just don't have enough information to make a judgment.

That's one suggested grouping. I don't know whether you may--and it doesn't mean that we need to fill in each of those cells. But that's the grouping that came to my mind as I reviewed this question when they were first sent.

Is that acceptable to the group? Okay. Then would you like me to call on each of you, or do you want a period of discussion before we do that? All right. Then why don't I start with Dr. Baker. Is that all right?

DR. BAKER: Let's see if I've got this charge properly. You want me to go through each of these measures and say whether they're mandatory for a growth study, whether they would be useful for a growth study, or not so useful, and then if they are experimental and not--

DR. GARZA: Without further research, you really can't put them in any box, but of interest, but, in fact, the research should be done.

DR. BAKER: Well, I think the first two should be mandatory without a doubt. The third, head circumference, it's certainly not in the experimental stage. I think it's accurate, but I don't see where it's useful for a growth study alone. So I would say it would be helpful but not mandatory.

I think the same thing goes for skinfold thicknesses. I think that the major amount of information would be--is obtained through the body weight and length. Skinfold thicknesses would be additional information that's useful. The accuracy is somewhat less, and I would say it's probably not mandatory.

Bioelectrical impedance, I don't think it's a stage where you can clinically use it, so scientifically I don't think it's appropriate yet.

Stable isotopes--

DR. GARZA: I'm sorry. Would you put that in the third category, No. 3, research?

DR. BAKER: Research. Stable isotopes I think are at the stage where you can use them but not practical, so I would put that as...

DR. GARZA: Can you define "practical" for the group?

DR. BAKER: It's not practical in terms of the money that it would require, the testing, the exposure, the instrumentation, the--

DR. GARZA: Is that relative to the benefit?

DR. BAKER: --the repeatability of the study, so I don't think it's practical to use it for a growth study.

The final one--

DR. GARZA: DXA is--

DR. BAKER: DXA. DXA I think is--would be in the category of very helpful but probably not required. Again, I think it relates to the feasibility of doing it, the availability of the testing, the time that's required, and, therefore, the money that's involved.

DR. GARZA: Any other physical body measurements or body composition measurements that are not in that list that you would like to add?

DR. BAKER: I'm not coming up with any. Underwater weighing--

[Laughter.]

DR. GARZA: Mothers may object. Okay. Dr. Stallings?

DR. STALLINGS: Body weight and recumbent linked, would those be essential? I have more interest in head circumference, and so I put that in the second category of moderate interest.

The two skinfolds that we discussed, triceps and sub-scapular, I would put in moderate interest. I think

bioelectrical impedance will not play out, so I have it in no use in the age group we're looking at.

Stable isotopes research, DXA in moderate interest because I think it may be a method that will help us in the future if we go down that way.

And the only other physical measurement that I felt like was missing would be the mid-arm circumference, and there's a history in neonatal care of using a head circumference, mid-arm circumference measurement that at times has been helpful in the research. So I would put that with the skinfolds under research interest.

DR. GARZA: Can we get that down? Dr. Heubi?

DR. HEUBI: Can I say "ditto"?

[Simultaneous conversation.]

DR. GARZA: --literally with everything, but--

DR. HEUBI: I don't have really much--

DR. GARZA: But who's ditto because they differed a little bit?

DR. HEUBI: Ginanne's.

DR. GARZA: Okay. I think I almost totally agree with what she said. I don't know that I'll add much by making any more comments than that.

DR. GARZA: That's fair. Dr. Anderson?

DR. ANDERSON: Obviously, this is not an area of expertise for me, but based on today's discussions--

DR. GARZA: This is for term infants. We're going to do this again for preterm, so I just want to make sure everybody understands that.

DR. ANDERSON: It seemed to me that my enthusiasm for head circumference is somewhat greater than what I've heard so far in light of what we heard about the relative correlation of length to head circumference and its ability to potentially identify deviations from standard neurologic development. Beyond that, I don't have anything to add.

DR. GARZA: You would classify head circumference then in the necessary--

DR. ANDERSON: Yes.

DR. DOWNER: Body weight, recumbent length, and head circumference would all be mandatory. Skinfold thickness I would put moderate. Bioelectrical impedance, do some more research on that. Stable isotope as well as the DXA, additional research. And for other physical body measurements, I too think that arm circumference may be of interest.

DR. GARZA: And you would put it in which category?

DR. DOWNER: Not moderate, but additional research. Additional research.

DR. GARZA: All right. Thank you.

Dr. Sigman-Grant?

DR. SIGMAN-GRANT: The first three essential, body weight, length, and head circumference. Skinfold thickness, I think it's moderate interest at this point. Bioelectrical impedance, I'd say no. Same thing for stable isotopes. It's a research, but it wouldn't be mandatory. I think there's potential in DXA, and I can't think of any other physical body measurements, but I think we should start thinking about getting out of the box, and maybe there could possibly be.

DR. GARZA: Any examples?

DR. SIGMAN-GRANT: Nothing that wouldn't be like MRI or--I don't even know, echograms or something like that. We can look at--but it's not my area particularly.

DR. GARZA: All right. Yes?

DR. MOYER-MILEUR: Moyer-Mileur. I would agree that the first three, body weight, recumbent length, and head circumference, should be mandatory. Skinfold thickness would be of moderate use. Bioelectrical impedance I don't think has a valid reason for use in infants. Stable isotopes for research, DXA I think would--could prove to be very useful with a number of caveats. And then there are other physical body measurements such as air displacement and TOBEC that potentially would be of moderate usefulness.

DR. GARZA: Okay. Dr. Kuzminski?

DR. KUZMINSKI: Sure. Thank you. Kuzminski. Again, this is not my area of expertise. I only know what I've read in the book and what I've heard in the discussions today. As mandatory, I would think body weight, recumbent length, and head circumference. Interesting, but not mandatory, certainly, skinfold thickness, but interesting.

I agree with the comment of looking towards the future, thinking out of the box in terms of traditional measurements, and maybe DXA falls into that category. And the others I would classify as research.

DR. GARZA: So you would put DXA into the research category or into a category that ought to be requested or--I'm--

DR. KUZMINSKI: Useful category.

DR. GARZA: Okay. Now, just to help the recording, it may be difficult. We've got, gee, it should absolutely be of use, moderate, might be useful, so you would put it in that second category then, DXA, along with skinfold thickness? Okay. Good. I wanted to make sure I understood that. Thank you.

DR. DENNE: Denne. I would also include body weight, length, and head circumference as mandatory, and I would put head circumference there because it really is our only surrogate measure of neurologic outcome and reflective of brain growth. So I think it belongs there.

Skinfold thickness is sort of moderate. Bioelectrical impedance, as discussed before, probably not useful, and the other techniques, research.

DR. GARZA: Okay.

DR. THUREEN: Thureen. I agree that the first three are essential. Skinfold thickness is of moderate interest, and I always do mid-arm circumference with skinfold, so I'd put those in the same group.

Bioelectrical impedance I think is probably of no use at this point.

Stable isotopes and other physical measurements are research tools, and DXA I have mixed feelings about because I think in most instances it's a research tool because it's not widely available. On the other hand, I think it's probably going to be of some significant use in the future, so it's kind of between the second and fourth category for me, but at this point I guess I'd call it a research tool.

DR. GARZA: Okay. Thank you.

Dr. Briley?

DR. BRILEY: Briley. The first three I think are of great use, skinfold thickness moderate, and the bioelectrical impedance, not yet. And the isotope and the DXA are research tools. I just wish that industry could get it down to it would be less costly piece of equipment. And the last is research also.

DR. GARZA: All right. Roger, would you have any comments?

DR. CLEMENS: I certainly agree with the [inaudible - off microphone]. I certainly agree with the rest of the [inaudible] research tool [inaudible] more cost-effective, technology will change [inaudible].

[Laughter.]

VOICES: Say it again.

DR. GARZA: No, we won't ask him to do that. That's all right.

There is a fair consensus on the first category in terms of body weight, recumbent length, and head circumference. I realize that it isn't unanimous, but generally as close as we come. Would anyone like to object to it being characterized in that way on the committee? With skinfold thickness as being of moderate interest, DXA at times was placed there, at other times it was put with research, but predominantly as a research tool. Is there any objection if the minutes record that in the end the discussion said, well, it's more of a research tool right now but one that we would recommend highly to FDA to get more information on because of its potential usefulness in future assessments? Or do you want to say no, it's of moderate interest? Ginanne?

DR. STALLINGS: I'd--

DR. GARZA: You need a mike. I'm sorry. Otherwise--I may get away with not asking you to repeat once, but possibly not twice.

DR. STALLINGS: Stallings. I'd like to pause and have a little bit more discussion of that, because I think it is going to be very available and is in most centers--and if you do some of the study kinds of things, like having the phantom and centralize the assessment and that sort of thing. I think it's moving forward fairly quickly. So I'm lobbying to keep it on the--out of the research only and into the moderate. Certainly it might not be appropriate for everything we're doing, but there's so much interest both in body fatness and the childhood issues related to osteoporosis that I think those two--this is the measurement that give you some assessment of bone health and some assessment of body adiposity. So that was why I was putting it in the moderate rather than research only.

There's a comment from the floor. I don't know if that's not possible at this point.

DR. GARZA: Not unless you want to ask them a question.

DR. STALLINGS: I don't.

[Laughter.]

DR. STALLINGS: I don't want to break the rules.

DR. GARZA: That's certainly within the procedure. You can ask if you want.

DR. STALLINGS: So that was why I put it in-- sort of started that.

DR. BRILEY: Could I ask you a question?

DR. GARZA: Yes, Dr. Briley, that is certainly within the procedure.

DR. BRILEY: What percent of the centers in the United States currently have access to this kind of equipment? This is not something I know about, so kind of fill me in about how many are already using it.

DR. STALLINGS: Well, it's like what we've been doing today. You'll have to define "centers."

DR. BRILEY: Whatever you--

DR. STALLINGS: But I think most children's hospitals are going to have them or have them. I think that Level 3 nurseries, which is where a lot of the sicker babies are often at adult hospitals, and they certainly have them. We will probably never see, you know, private pediatric practices having them, but they would have access to them in their community, generally under the auspices of women's health centers and things like that. But you've heard all the caveats about if you really are going to do this, you have to have a technician, an operator who's really good with kids and moms, and so they're all things that flow from that.

But the current DXA technology is fairly widely available in middle-size towns and up and in all major medical centers.

DR. GARZA: Virginia, related to that, what percentage of centers that have an interest in doing growth or nutritional studies to you estimate now have DXA? Is that more than half or less? In terms of groups that might be engaged in this type of assessment.

DR. STALLINGS: Well, I think the groups that really define themselves as interested in growth, both in little--you know, failure to thrive and obesity, are going--either have it or are approaching 100 percent, because it is--it was a breakthrough technology in childhood body composition and growth studies. So I think if they don't have them, it will be hard to stay in the field, and most of us who have gone from a lot of stable isotope work to this find this a much easier technology to work with with children of all ages.

DR. GARZA: Dr. Heubi?

DR. HEUBI: My comment was, although this is widely available, not everybody has pediatric software, and that becomes a bit of an issue. But Ginanne is absolutely--she's absolutely correct about it. Most centers that have GCRCs that are in pediatric centers have them, and they're broadly available. And they're often free and at low cost in the GCRC--

[Laughter.]

DR. GARZA: That sounds like a paid political announcement.

DR. HEUBI: This is an infomercial.

DR. GARZA: An infomercial, that's right.

DR. STALLINGS: He's the center director.

[Laughter.]

DR. HEUBI: But it's true that most of them have--actually, I'd say--there are eight pediatric centers in the country, and I'd say pretty uniformly they all have them. So it's widely available.

DR. GARZA: Would anyone like to speak against Ginanne's proposition?

DR. DENNE: I guess I'd just want to raise the question about what are you going to do with the data. You know, how are we going to interpret the DXA scans? Again, I'd love to have that as a reference, but how are we going to, in the context of a formula study, interpret the results?

DR. STALLINGS: Well, you know, we all can see that we don't have great reference data right now. And, in fact, there's a major NIH-funded study looking at reference data down to a certain age that's going to address that in older children. And I think one of the gaps that remains is having similar quality--I think it's a five-center study? I think one of the challenges that

remains is the reference data, but I would propose this would be used in the randomized trial setting where you would have enrolled groups randomized, and then you really could start to see are we having differences.

My interests really are in all three compartments, actually, that--are we making children fat to the detriment of the fat-free mass? Or are we making them fat and fat-free mass is fine? And there are a number of issues that are coming along in the infant formula world and in the antecedents of osteoporosis that really directly have to do with bone mineral accretion.

Now, it would be completely misrepresenting it to say that what goes on at 4 months we know has something to do even with 2 years being now ambulatory or at 7 or at that critical time sort of 10 to 14. So--but, I mean, it is, I guess from the committee point of view, you know, it's futuristic, but I'm already pleased that we've added length to this. You know, so we are trying to think about where do we want to go in the future.

So, to me, the big deal is we've changed from just worrying about children not growing, which is failure to thrive, to trying to look at both sides of the growth spectrum.

DR. DENNE: And, again, I don't have--I think that data would be useful, but specifically if you find a

difference in one compartment or another, what is it that you will do with that between formula?

DR. HEUBI: I think what it's going to require is the more longitudinal view, because one of the comments was made about not knowing anything below age 2 in terms of what its relationship to adult obesity is. This would give us an opportunity to actually go another step further with follow-up studies in the future. Again, it would require probably out of the scale of an infant formula trial but would be some information that could be added as added knowledge.

DR. THUREEN: Thureen. I agree with what you said, the usefulness as a tool. I just don't want to--this committee's recommendation come across as meaning that unless you have a DXA you shouldn't be involved in formula trials. And I'd hate to see that happen because I know at most centers, pediatric GCRCs are at most children's hospitals, but there are still a lot of people out there who don't who have been involved in formula trials. So I think we have to make it very clear that this--when we say moderately useful, it shouldn't be looked on as that kind of a center would have a preference for doing these types of studies if they had a DXA machine, because I don't think at this point in time that's realistic.

DR. GARZA: Okay. There is still is somewhat of a split, is my sense, but if it's not required, then I think we've given the FDA a sufficient breadth of views on this that prolonging the discussion probably will not be very useful. I don't think that there is a fundamental difference in how this technique is viewed, but where it might be placed.

Okay. Bioelectrical impedance was pretty much put into the useless category for right now, and with stable isotopes also being in the research category, and arm circumference possibly being placed alongside with skinfold thicknesses as of moderate use. Is that...any objections to that summary?

[No response.]

DR. GARZA: All right. Then we're done with terms. We'll start with preterms, and I'm trying to find...

Dr. Thureen, would you like to start?

DR. THUREEN: Yes, I will start. I think that body weight, recumbent length, and head circumference are critical. I think that skinfold thickness and DXA are of moderate interest, but are of less use at this point in time than they actually are in terms infants because of the technical difficulties in using them. And I think there is no role for bioelectrical impedance. I think that stable isotope and other physical body measurements

are not indicated at this time, are really in the research area.

I do want to stress that in this age group, I think head circumference is critical because we use head circumference frequently for detecting significant abnormalities in growth, and it is a major way of looking at--or at least suggesting long-term neurological growth.

DR. GARZA: Okay. Thank you.

Dr. Denne?

DR. DENNE: I don't think I would add anything different to that. That seems a reasonable position.

DR. GARZA: Okay. Dr. Kuzminski?

DR. KUZMINSKI: I have to defer to the experts on this and agree to the same thing.

DR. GARZA: Dr. Moyer-Mileur?

DR. MOYER-MILEUR: I would just concur for preterm babies that head circumference is probably more critical than recumbent length and easier to obtain.

DR. GARZA: I didn't sense in the previous discussion--just to make sure that the minutes reflect the group's sentiments correctly--that head circumference, recumbent length and body weight were prioritized. We said all three are required at the same time.

DR. SIGMAN-GRANT: I agree with everyone else so far.

DR. GARZA: Sigman-Grant.

DR. DOWNER: Downer. Ditto for me, too.

DR. GARZA: And that's Dr. Downer.

DR. ANDERSON: Anderson. The same.

DR. GARZA: We're on a roll.

DR. HEUBI: Heubi, except that I do want to say I'm glad that Patti came around to thinking that DXA was of moderate importance.

[Laughter.]

DR. STALLINGS: And I would agree, and I was one of the ones who didn't put head circumference in priority for term, and I agree completely it's essential for preterm. And I was just trying to have some distinguishing characteristics.

DR. GARZA: Okay. Thank you.

DR. BAKER: I think I agree with everyone. I would vote for head circumference being essential for preterm growth studies.

DR. GARZA: All right. Dr. Briley?

DR. BRILEY: I agree with Patti.

DR. GARZA: All right. Thank you.

Roger, would you like to make any comments?

DR. CLEMENS: I certainly concur with the group.

DR. GARZA: Okay. Well, gee, all right.

[Inaudible comment off microphone.]

DR. GARZA: You can repeat it this time.

DR. CLEMENS: Officially, yes, I concur with the group.

DR. GARZA: You have to identify yourself.

DR. CLEMENS: Roger Clemens. I'll get it out yet.

DR. ANDERSON: Then I think we've answered Question 2 in the way the groupings were made. Does anyone want to address any aspect of two that you don't think we've addressed?

DR. ANDERSON: On behalf of the group, I say no.

[Laughter.]

DR. ANDERSON: This is Jim Anderson.

DR. GARZA: Thank you, Dr. Anderson.

All right. Then let's move on to No. 3, and 3A is that the metrics above can be evaluated as either retained or absolute growth or velocity, rate of change. Comment on the distinguishing values and merits of each static or variable method in the assessment of normal physical growth. I'm going to limit the discussion to body weight, recumbent length, head circumference, skinfold thickness, because, in fact, none of the others were either recommended or seen as useful, so I don't see much point in our discussing each of those.

Who would like to start? Dr. Stallings?

DR. STALLINGS: I think that both attained growth and velocity are essential for the first three,

for weight, length, and head circumference. I think velocity would not be helpful in skinfold thickness because of the very small magnitude that we're moving through. And in the age range we're looking at, I honestly don't know about mid-arm circumference since we've clustered that one. I think the change is relatively modest, but I'll defer on that one.

I think that covers it.

DR. GARZA: Well, the only other thing that might be useful, although it may be frosting on the proverbial cake, is whether you feel that velocity, since you said that it would be helpful for body weight, recumbent length, and head circumference, should all be obtained with the same frequency or whether you want to suggest any frequency of measure to be able to accurately reflect velocity.

DR. STALLINGS: Well, there are two issues here. I'm still of the mind that most of the things that merit growth study would have a contemporary, comparable group, so that I would be measuring them--if you were going to existing data, I think the velocity curves are 1 months at this--birth to 4 months?

DR. GARZA: I just meant in terms of how often would you want measures of body weight, recumbent length, or head circumference obtained to be able to define velocity.

DR. STALLINGS: In a study--

DR. GARZA: In a study population. Are three measurements sufficient, or do you want more frequent measurements?

DR. STALLINGS: A baseline and--well, baseline, and then assuming that was before 1 month, because everything we've heard said that that should be, and then 1 month, 2, 3, and 6. Again, I guess I don't understand how long we're planning to run this. A 4-month study?

DR. GARZA: Four to six months would be my--

DR. STALLINGS: Okay. So baseline, 1 month, 2 months, 3, 4, and 6.

DR. GARZA: I'm just basing that on the previous question that said birth to 6 months.

DR. STALLINGS: Right. Good.

DR. GARZA: I'm sorry. Could you repeat that again? You said birth--

DR. STALLINGS: Baseline--well, birth, the 14-day, then 1 months, 2 months, 3 months, 4, and 6.

DR. GARZA: So, in essence, monthly for the first 4 months, and then 6 months, the first month having measures at birth and at 2 weeks.

DR. STALLINGS: Right, or whatever that earliest--but something--I was calling it baseline.

DR. GARZA: Between 1 month and birth.

DR. STALLINGS: Right. But something before--at 14 days as your baseline, and then again at 1 month.

DR. GARZA: Okay.

DR. STALLINGS: So that might be a 28-day or a 14-day measure.

DR. GARZA: I don't know that we have to be that exact, but it gives people a sense of how velocity would be obtained, or measured, at any rate.

Dr. Heubi?

DR. THUREEN: Dr. Garza, can I make a comment?

DR. GARZA: Sure.

DR. THUREEN: I'd like to ask the committee if they'd agree on the time points that Dr. Fomon suggested this morning, the seven time points over the first 6 months of life, because I think that's what you're referring to. And unless there's dissension from that, which I suspect there's not but I'd like to ask the committee, that those may be the time points that we'd recommend for longitudinal studies.

DR. GARZA: They pretty much coincided with those.

DR. THUREEN: But we could just call it the Fomon criteria from now on rather than--

DR. GARZA: That sounds good. Thank you, Patti.

Dr. Heubi?

DR. HEUBI: If you'll give me one moment, I just want to see what he actually said.

DR. DENNE: This is Denne. I think it was Dr. Chumlea who suggested that time frame. I think Dr. Fomon was sort of talking about 4 months as the criteria.

DR. HEUBI: I think birth, 2 weeks, 1, 2, 4, 5, and 6 months.

DR. GARZA: So the third month was omitted. I think that's--we just need a general idea. I don't know that we have to come up with the exact metrics or exact time points. That's close enough. There'd be some wobble.

DR. SIGMAN-GRANT: This is Sigman-Grant. Is there any advantage to having a measure between birth and 2 weeks?

DR. GARZA: Between birth and 2 weeks, is there any advantage? There is quite a bit of wobble, is the sense that I'm getting from the group again at the table because of weight loss during--

DR. SIGMAN-GRANT: But is that an important measure?

DR. GARZA: There are no standards that I'm aware of.

DR. STALLINGS: Right. I think what you would end up there mostly would be with the individual variation and hydration and immediate pre-partum kinds of

things, which wouldn't reflect the feeding experience as much.

DR. SIGMAN-GRANT: Okay.

DR. STALLINGS: And, also, just--I mean, I don't want to talk about practicality, but, you know, there is the issue of if we could get a birth, true birth measurement, and then 14 days or something like that, that would be great. But I think what we're doing is just skipping the variability of when you regain your birth weight and establish full feeding.

DR. GARZA: In the interest of time, while Dr. Heubi's getting these things together, would you--maybe we can take both preterm and term together? Would you change any of those recommendations for the preterm, or would you leave them pretty much the same for both term and preterm?

DR. STALLINGS: I think I would leave them pretty much the same, again, just recognizing the preterm--that the physical measurements on the preterm are more stressful, but we still have to have them to be able to make the decisions we need.

DR. GARZA: Would you recommend--well, going to 6 months post-conceptional equivalency or-

DR. STALLINGS: Yes, I would.

DR. BAKER: I think that both the static and the velocity measurements should be obtained for weight,

length, and head circumference. I don't think velocity is necessary for skinfold thickness or mid-arm circumference. And I would agree with the time points as Ginanne said them, and I also think that it would be useful for both full-term and premature babies.

DR. GARZA: Okay. Thank you.

Dr. Heubi?

DR. HEUBI: Well, after all that, I don't think that I disagree.

[Laughter.]

DR. HEUBI: I was looking very feverishly for this information. I think Jim showed it to me. So I don't think that--I think that I would agree with the same time points and the same measurement parameters would be quite appropriate, and for preterm infants, I think it would be 6 months post-conceptional age. I think that's the issue that has to be addressed, and I think that should be left up to Dr. Denne and Dr. Thureen.

DR. GARZA: All right.

DR. ANDERSON: Anderson. Agree.

DR. DOWNER: Downer. I, too, agree. Body weight, recumbent length, and head circumference at birth, 2 weeks, 1 months, 2, 3, 4, and 6 month intervals. I'd also agree that skinfold thickness and mid-arm

circumference, the velocity for that is not important here.

DR. SIGMAN-GRANT: Sigman-Grant. Ditto.

DR. MOYER-MILEUR: Moyer-Mileur. I agree for the term infant on the attained and velocity measures as well as the measurement periods. But for the preterm baby, I think we need to keep in mind that some of these studies will be done in-hospital versus post-discharge, and I think in the hospital that your intervals need to be different. And so it would be wherever--a set baseline, and then probably every two weeks until that child is discharged, and then go into the 1-month interval to 6 months post-conception age.

DR. GARZA: Okay. Dr. Kuzminski?

DR. KUZMINSKI: Thank you. I agree.

DR. DENNE: Yes, I would agree for the--

DR. GARZA: Dr. Denne?

DR. DENNE: I'm sorry. Yes, I would agree for the term, and also agree that preterm needs more frequent measurements in-hospital, probably every 2 weeks, every week to 2 weeks, something like that.

DR. GARZA: Dr. Thureen?

DR. THUREEN: Thureen. I agree on the term infant, and, again, the preterm infant, it's going to be really critical looking at what their growth rate is over their hospitalization period, especially when they're

sick. So initially it may need to be every week, but probably no less than every 2 weeks during hospitalization.

DR. BRILEY: And I agree.

DR. GARZA: I'm going to ask the group the question of whether somebody is willing to comment on the merits and value of having both static and velocity measurements. I think we've got fairly good agreement, other than on the number of measurements, and the frequency of measurements of premature. Would there be any objection if we summarize it? In fact, the in-hospital phase of the premature management should probably be more frequent. Every two weeks seems appropriate.

DR. STALLINGS: That was actually part of my question because the--Stallings. I think the question you bring up about weekly measurements, while hospitalized--

DR. GARZA: She said every two weeks, I think.

DR. STALLINGS: No, it was--

DR. GARZA: Is it weekly or--weekly.

DR. STALLINGS: Both neonatologists considered weekly measurements, so that was why I wanted to bring that back up because I would agree with that, if you think it's merited, because things are happening so fast, and they could either be in a rapid rate of growth or a

very clear static phase. So I was asking the two neonatologists to comment on one versus two weeks, and then the measurement error and all of the things that go from that.

DR. THUREEN: This is Thureen. I would advocate for one week because the fastest rate of growth is actually the in utero rate of growth, and so potentially, if these infants are growing at the in utero rate, you've got it higher than you have during the growth spurt post-natally? So I think that there are going to be dramatic changes even over a week period, and I would advocate for weekly measurements.

DR. DENNE: Yes, I would basically agree. It depends a little bit on the study design, and when patients get entered and those kinds of things. I guess the principle would be more weekly measurements early and spacing out later, but--

DR. GARZA: Dr. Anderson?

DR. ANDERSON: I would only ask that we consider the distinction between what's required, for lack of a better term, good patient care, and what's required for the purposes of documenting normal infant growth in the setting of an infant formula study. I wonder whether, in the latter, every two weeks might not be sufficient.

DR. THUREEN: Thureen. I think pretty much every neonatal unit measures these parameters on a weekly

basis anyway. Whether the actual research measurement would be more destabilizing I think is difficult to say. I'd also say that in most units, if your patient is unstable, you don't even do your routine measurements. You just have to have some latitude for those kind of decisions, so ideally every week, but if the patient is not stable or if it appears to be a difficult measurement for a patient, then I think you may have to be a little bit more flexible.

DR. DENNE: Yes, and obviously that partly depends on the study design, but things do happen quickly over a week. These measurements are clearly possible over a week, as opposed to later on when it's tough to make that difference over a week.

DR. GARZA: Roger?

DR. CLEMENS: I do want to refer to that. In terms of body weight and length, we do, the infant formula manufacturers, do conduct these studies, and when they do, they measure length and weight at the 14, 28, 56, 84, 112 days. Clearly, the Fomon formula, if you will, when we conduct four-month studies, and certainly to include two more data points would not be a big deal if required by statute to go on to six months for these kinds of growth studies.

We do collect skinfolds, but not as frequently at this point in time.

DR. GARZA: Any other comments regarding the one or two weeks for the prematures?

[No response.]

DR. GARZA: All right. What about the values and merits of attained and velocity, does anyone want to speak to those any more than you may have in your comments?

Dr. Anderson?

DR. ANDERSON: Perhaps I'm not thinking of this correctly, but it does seem to me, if we have the data points at the specified times that we've talked about, then the data can be presented, in either way, simply through a data transformation. And so if people are normally used to thinking about velocities, they can be recorded as velocities, and if they're used to thinking about weights and specific I points, they can be recorded that way.

DR. GARZA: Any other points that anyone wishes to make in response to that request from the Agency?

[No response.]

DR. GARZA: Then why don't we move on to 3B, and here again I think we can take both the preterm and the term together. The outcomes above can also be evaluated as individual infant data or as group comparative data, comment on the values and merits of using individual or

aggregate data in the assessment of normal physical growth.

DR. DENNE: Dr. Garza, before we start, I'm not sure I understand that question.

DR. GARZA: I think what they're asking is should the Agency be presented with the individual growth data for individual subjects or just the aggregate data for the entire group.

DR. DENNE: So in terms of what's actually submitted to the Agency.

DR. GARZA: Or here it's phrased, in terms of the values and merits of each one, but I suspect that that's why they'd like to know that.

Is there value in submitting both and or merit, and would one or the other suffice?

Dr. Anderson?

DR. ANDERSON: I think that either would suffice. For longitudinal data like this, there's often a choice in terms of how one both presents and analyzes such data, and it reflects essentially the question here; that is, one can analyze parameters at the patient-specific level and then estimate population parameters, considering the patients as random realizations from that population, so each patient would have specific parameters, some showing a strong growth tendency, others less so, and that the parameters, the estimates would be

the average of those parameters in the population or one could take what's often called a population-average approach, which is essentially one is not interested in the parameters of individual subjects, but how they're realized in the observed population, and there the focus is more naturally on the population values that are observed at individual time points.

My own personal view is that the answer that you get from the two approaches are oftentimes largely the same, and that I would think that either approach here would be acceptable.

DR. GARZA: But what sort of information would you recommend be obtained if population or group data are being submitted that would describe the distribution? Would a mean and standard deviation be sufficient or do you think we ought to look at "skewdness" or whether, in fact, it's normally distributed? Can you elaborate a bit on what you think would be the essential information you would need to assess the distribution?

DR. ANDERSON: Well, we're largely getting ahead of ourselves because we haven't talked about what we would be comparing things to.

DR. GARZA: Assuming only one group, and you're going to look at individual data versus group data.

DR. ANDERSON: If it was a single group, I would think that you would want more than simply means and

standard deviations because, as we've seen from the, from the growth charts, the data can be somewhat skewed at various time points, and so either some summary of the data reflects a transformation towards normality at each of the time points or some other type of display which allowed one to see not only the measure of central tendency, but also a sense of the general distribution at the times when the data were collected would be appropriate.

DR. GARZA: Drs. Heubi or Downer, would either one of you like to go next?

It looks like you're the winner, Dr. Heubi.

DR. HEUBI: Did I win? And it's still on.

I tend to think that it probably would be appropriate to submit all of the data showing individual data points, in addition to some summary data, just because it looks to me like based upon the size of these groups are likely to be such that it would be a value to see what the distribution looks like, in addition to the summary statistics.

DR. GARZA: Can you elaborate a bit on the value of seeing both?

DR. HEUBI: It would allow someone to look at how normally distributed the data was and how whether there were outliers that were not reflected by the summary statistics.

DR. GARZA: Dr. Stallings?

DR. STALLINGS: Stallings. I would like to see both. The reason for the individual data I think is important to, I mean, I can't give us a good example of what's happened historically, but I think it would be important to see a group of children that would be on both ends of the distribution, and again particularly also by gender. I mean, I think there are some different things going on, so I'm assuming we get the different segmental analyses that we had talked about.

So I would be really interested in seeing how many individual babies fell out less than a fifth, less than a third greater than the 95th. Now, the logical question is exactly what am I going to do with that data, and I don't really know that, except that I think if I were in a regulatory environment and I saw that, yes, most of the kids were doing just fine, but there really was a group that wasn't, I would want to explore that more, and that might start the dialogue to find out more about that.

DR. GARZA: Dr. Baker?

DR. BAKER: I would certainly want the individual data I think, just thinking about the same sort of things that there are a possibility of outliers that would trigger you to think more about the whole problem. I don't think we can give guidelines about what

exactly you would do, but I certainly think it's worth looking at.

I think some group data would also be useful sort of, but I think certainly the individual data and a good bit of the group data would be useful.

DR. GARZA: Dr. Clemens?

DR. CLEMENS: Clemens. You caught me early.

DR. GARZA: I called you earlier, that's right.

DR. CLEMENS: I appreciate that. Actually, the pro forma to manufacturers provide all of the data, to provide individual data, as well as group data, and so they can clearly see, to your comment there, Ginanne, they can determine if there's any outliers, if you will. Also, if there are any outliers, potential outliers, as there is in any population group, then, actually, the clinical or the medical director provides clinical comments, as well as to the principal investigator, to provide clinical comments specific to that particular child, so you get here exactly each child.

DR. GARZA: Thank you.

Dr. Downer?

DR. DOWNER: Thank you. I'm still pondering this. I think both data sets would be important. As Dr. Stallings said, I'm not sure what we would do with the individual data, from a group perspective, but it would provide some information on the substance within the

group, and that's important. So I think both data sets would be important.

DR. SIGMAN-GRANT: Sigman-Grant. I think both sets are important. I think, if you look at individuals, you might see clustering, perhaps by ethnicity or something else that you might not see if you only saw population data, and that might be helpful information. So that's like clustering of individuals.

DR. GARZA: Dr. Moyer-Mileur?

DR. MOYER-MILEUR: Moyer-Mileur. I would agree that both individual and summative data is important.

DR. GARZA: Dr. Kuzminski?

DR. KUZMINSKI: Both data, please.

DR. DENNE: Denne. I would concur with the previous comments.

DR. GARZA: Both in terms of the--

DR. DENNE: Yes.

DR. THUREEN: Thureen. I would concur. I have nothing to add to the comments that have been made from that side of the table.

DR. BRILEY: Briley. I would concur. I also think about this might be the first time that we have these data, and the next time around, it would be nice to do some comparison with it. So it would be nice to have both.

DR. GARZA: All right. Does anyone want to add any comments to what you may have said earlier on this question?

[No response.]

DR. GARZA: If not, there seems to be a consensus that both individual and group data should be evaluated, and the values and merits, having to do with the ability to assess the distribution and potential outliers, et cetera, in a way that perhaps summary statistics may not lend themselves to easily.

It is 5:30, almost. We can go ahead and try to do the next one or two questions, if you're up to it, or we can try to get the van here by 6 o'clock, and we will still be here. Let me confer a bit. It'll be here at 6:00, so why don't we go ahead and do the next one.

DR. SIGMAN-GRANT: It's a really big one.

DR. GARZA: It is a big one, you're right.

Dr. Briley?

DR. BRILEY: Don't you think we can do the question now and not have to do the work tonight? Would that be a trade-off you'd allow?

[Laughter.]

DR. GARZA: Oh, that doesn't buy us more time, Dr. Briley.

DR. SIGMAN-GRANT: This is Sigman-Grant. I think both of these require a lot of, it's going to be a

lot of discussion, the difference between reference and standards.

DR. GARZA: You notice that I said it with some trepidation, and you're absolutely right. I'm trying to think how one could use the next 30 minutes most productively, given the fact that the idle mind tends to be the devil's workshop.

Why don't we then work on Question 7 on your own, as Dr. Briley suggested.

Dr. Denne, did you--

DR. DENNE: Well, I just wonder whether Question 4, whether there wouldn't be reasonable consensus about the answer to that question. I mean, my sense is there might be, but--

DR. GARZA: We have one potential masochist in the group.

[Laughter.]

DR. GARZA: Do I see another one that's willing to start with No. 4?

DR. STALLINGS: Take Question 5.

DR. GARZA: Well, 5 might benefit from having 4 in the discussion, and that's why I didn't want to take them out of order. I thought of that, and so why don't we begin and then see whether, in fact, we are making progress, and we don't have to come to a forced conclusion because Dr. Sigman-Grant is right, it's

probably going to take quite a bit of discussion, but Dr. Denne may prove us wrong, in the fact that we can come to a consensus in 30 minutes, and that would save us an enormous amount of time tomorrow.

So let's, since the bus can't pick us up until 6:00, let's try and do 4:00. It's for adequate evaluation of normal physical growth. Below are examples of clinically distinct reference groups, and we really have three. One, our concurrent controls, a reference data uses a control, and historical control, and I'm assuming that by reference, what is meant is a more comprehensive data base with historical controls, perhaps referring to a specific formulation that might have been looked at in the past, but not necessarily information that may be in the public domain. Is that an accurate interpretation of that? Dr. Walker is saying yes.

So a reference data, think of it as something that is in the public domain that is quite extensively documented, as opposed to a historical control that might be more limiting in that it may just be either proprietary information or information that relates to a specific formulation from a specific one single article or one single study.

So it might be seen as a historical control, but not necessarily as a reference. I'm at a loss as to what other groups, but let's get those on the table.

DR. HEUBI: I guess it's possible that that is referenced as longitudinal data, as opposed to other, which might be cross-sectional. That would be another option for other.

DR. GARZA: You mean that would be under reference or historical that it could be either longitudinal, the substance of those?

DR. HEUBI: Well, I'm saying reference data used as controls would be things like the Fels-Iowa data, as opposed to the CDC data for other as cross-sectional.

DR. GARZA: I'm sorry. In my description, I meant to include both Fels and CDC, for example, as reference data.

DR. STALLINGS: So not just longitudinal--

DR. GARZA: And subsets of that. It could be cross-sectional.

DR. HEUBI: it would be relatively easy to rank these if you separate those out and make it the other way, so it's cross-sectional being your last category.

DR. GARZA: Okay.

DR. HEUBI: Because that's the least desirable.

DR. GARZA: So you would say cross-sectional, and we would interpret concurrent controls, reference data, and historical as longitudinal, and then other would be cross-sectional or would you prefer seeing six

cells with concurrent controls being longitudinal or cross-sectional? No?

DR. HEUBI: This may be a long time.

DR. GARZA: That's more complicated, but I think it's going to be complicated by just saying all of the above is cross-sectional, but let's try it.

Yes?

DR. SIGMAN-GRANT: I have a question.

DR. GARZA: Sure.

DR. SIGMAN-GRANT: Sigman-Grant. Are we referring to this as reference or standard?

DR. ANDERSON: This is Anderson. I think we can't do anything other than refer to this as reference because that's what they are.

DR. GARZA: So that the implication of that is that you're not going to be making any value judgment on the basis of the comparison, but that you're suggesting we limit the discussion to just the fact that if it's going to be a comparison, it'll be more or less, not better or worse. Is that clear to everyone? Because it's an important distinction.

Now, we may want to take up the issue of standard in number five, but for right now we would limit this to a reference, and so that what one would say is more or less, not better or worse.

Dr. Anderson?

DR. ANDERSON: With the permission of the chair, let me take a stab at this.

DR. GARZA: Yes, please do.

DR. ANDERSON: My own personal view is that the least helpful of what we've discussed are historical controls, largely because they tend to be small populations collected in a process which is often difficult to describe, and the temporal changes that can occur make it highly likely that historical controls will--it will be extraordinarily difficult to convince individuals that the historical control accurately reflects an appropriate comparison group.

And then notwithstanding my attempting to push the envelope during the discussion today, I think that using reference data is not that much better, largely because it's just that. It's reference data, and so as a result of the conduct of a study, we can say that the results were greater than or less than the standard--sorry--the reference, but I don't know how to interpret that. So we end up with the study group having a mean weight at six months, which is the 55th percentile of those standard.

The conclusion of that is I don't know how you get beyond that, and so my view is that if it's felt that it's important to evaluate normal physical growth, it

should be done in the context of concurrent randomized controls.

DR. GARZA: Dr. Clemens?

DR. CLEMENS: Clemens. I trust that we don't make a decision recommendation this evening on this, Bert. Tomorrow, we have some public comment available and also the three medical directors for the manufacturer was making comment on this question, and other related questions, which we have not addressed tonight. So, again, I trust that we don't make a recommendation to the FDA this evening.

DR. GARZA: We can hold off on that until tomorrow.

DR. CLEMENS: I would trust that--discussion is good, but no recommendation at this time, please.

DR. THUREEN: Thureen. I'd like to note that Dr. Anderson's comments closely reflect Dr. Frongillo's report of earlier today, where he really elegantly laid out the advantages and disadvantages of concurrent reference and historical reference data or controls, and I was already convinced this morning, after his talk, and I'm now more convinced after hearing your opinion.

DR. GARZA: One of the things that I'm reading into what you said, Dr. Anderson, is that the absence of standards really limits us in terms of making those comparisons, so that having a concurrent randomized

control group makes a lot of sense; is that one of the major factors in your conclusion?

DR. ANDERSON: Yes, you said that very well. I think that if there were data which were generally agreed reflected in an agreed upon standard, and I have not heard that today, then the discussion could be broadened because then it would be possible, for instance, to say that any outcome which led to results which were within .2 standard deviations of the standard would be considered acceptable evidence of normal physical growth, in the absence of a standard, and with the presence of what I gather are generally agreed upon as references, that kind of argument is not permitted.

DR. GARZA: Dr. Baker?

DR. BAKER: I'd just like to put in my two cents here. I think that we all agree that the gold standard, the thing we all like to see, is a concurrent randomized control group.

DR. GARZA: Is the mike on? I can't tell.

DR. BAKER: What we'd like to see is a randomized control group, and so that we put as number one, and then I would put a longitudinal reference group as kind of the second choice, and then a cross-sectional reference group as a third choice, and most often, the historical control is way down there.

On the other hand, I would have to say that there might be some circumstances where the historic control group might be reasonable, and that would be in a situation where you were doing rapid sequence studies with very similar formulas and just changing one thing. So you'd use, for those series of studies, you might use one historic control group, but--

DR. GARZA: You may want to hold the mike closer to you.

DR. BAKER: --but to think about using a control group from 10 years ago, seems to me, unreasonable.

DR. GARZA: Thank you.

Dr. Heubi?

DR. HEUBI: Heubi. I would agree with Rob. I think that historical controls are probably only acceptable if they're fairly recent, and whatever fairly recent means, I don't know, within a year, two, three years, but not ten years. But the ordering, I think I would agree with him totally, concurrent controls should be the gold standard, and a longitudinal reference group, like Fels-Iowa, would be a second, and then a cross-sectional and then finally historical controls.

DR. GARZA: Dr. Stallings?

DR. STALLINGS: I think the concurrent control group is essential. I think often when the rest of us are doing things that are research, we wouldn't be able

to do research without an adequate control group, and we're not setting policy or regulating anything or making products that, you know, millions of infants will be exposed to. So I really do feel strongly about that, and I concede that there may be a time, if you really are designing things, that the concurrent control group might be able to serve for a series off studies.

I feel like the historical, as defined when we started this, the historical group, which might be proprietary and might be specific to a single formulation, a small population, less-well described, I really don't think that's adequate sort of for the modern times, partly because I think it's not in the public domain.

And if we're saying you don't have to do the study, which might remain proprietary, that somebody really needs to agree that, other than the regulatory group, agree that it was a good choice.

So I think it does come down to the control groups, if we reach the level where a growth study is required, then it should be an optimal growth study.

The issue of the reference data, I really learned a lot over the last couple of days reading about this and listening to the comments. I think I have come to believe that because the reference data represent so many years, and we really don't know what's optimal, and

that's, you know, we need many more days conference on that, that I'm less inclined to use reference data. I would have it available for descriptors because what you would want to know, if there is some complete quirk, and now the control group and the study group happen to all be at the one standard deviation of everything, you would want to know that, that that's the data you would be looking at, but I don't think it would be--and I would worry about how we got into that fix--but that would be a bit different.

So I'm really voting for the contemporary control groups.

DR. DOWNER: Downer. I, too, agree that the historical control would not have much implications for the growth study at this time and that the gold choice, of course, the concurrent randomized trials, would be the one to select. The longitudinal reference and cross-sectional reference, too, would not be of great importance here, but it would be good to look at over the long haul just to see what has happened and to make comparisons over time.

DR. SIGMAN-GRANT: This is Sigman-Grant. I, too, agree that the concurrent controls would be the most ideal.

I'm sitting here reflecting. It's interesting because when I work prenatally, there's the issue of what

happens prenatally determines potentially what might happen postnatally, the so-called metabolic imprinting in Barker, and yet we seem to think that, okay, we're just looking from birth on, and so without a concurrent group and using previous information makes me question if, indeed, metabolic imprinting, as a hypothesis, is explored and expanded.

There's a lot of difference, maternal differences, between, say, the '50s, '60s and '70s and what's now recommended--weight gain, smoking, alcohol use. If that, indeed, affects infant growth, postnatally, which it probably does, it might, then if we use data that was presented for those years, are the infants truly representative of the cohort from today?

DR. GARZA: Are there any other comments that anyone would like to add? And we'll come back to this issue tomorrow after the public comment period. Any other--I'm sorry, go ahead. I didn't see you.

DR. MOYER-MILEUR: That's okay. Moyer-Mileur. I would just like to say I assume we're all talking from the term baby point of view at the moment, but for preterm babies, there is no reference. So I think for preterm studies they always have to have a concurrent control because medical care in the NICUs is changing constantly, and you can't really use a historical control from even a year in the past.

DR. GARZA: That's a good point.

Any other comments? Your free to speak today. We're going to just come back to it tomorrow. I want to make sure that we've got the time. So you can speak today or tomorrow.

DR. KUZMINSKI: Just a comment. I concur with the comments made by the rest of the committee on concurrent controls, as the gold standard, but I feel that there is place for historical controls. To me, it depends upon how, an agreement with another comment that's been made, how long back in time that history is, in terms of that data. I think a driving point also is how relevant was that study population to the current study population that is being proposed in the current study?

It gets back to my question, I guess, I posed to two of the speakers, Dr. Denton and Dr. Bier today, what is that gray area that we use experience to make a judgment not to use, not to do a clinical growth study, and do we just automatically default to doing a study? I think that's a very difficult question to answer.

This hits on it, also, on historical controls, and the reference data, I agree with the others as the third most--

DR. DENNE: Denne. I guess I don't have much to add. I would agree with the previous comments about the

value of concurrent controls and the significant lower value of any other approach.

DR. THUREEN: Thureen. I agree.

DR. BRILEY: Briley. I would like to make a statement in regard to what our current consumer is, in terms of it's different than what we've had in previous years, like Maureen had said, and so I feel like the gold standard is the concurrent controls. I think it has to be there because society has changed, and the young mothers today have a different role to play than what they did 10 years ago, and 10 years is too long.

DR. GARZA: Would any of you have any questions to any of the presenters that might have addressed this topic earlier today?

[No response.]

DR. GARZA: Then, we pretty much have dealt with, A, for term infants, we had Dr. Moyer-Mileur addressed the preterm as being a special case where, in fact, that's even truer in terms of reliance on a concurrent control.

Does anyone want to speak to preterms or would the discussion that we've had, in your view, pretty much reflect your sense of preterms as well?

[No response.]

DR. GARZA: People are pretty tired, Dr. Sigman-Grant is telling me.

DR. STALLINGS: Dr. Stallings.

Just to clarify, though, I think what Laurie was addressing, we had some diversity about whether concurrent controls were almost the only thing or they are second and third rank. I think you were suggesting, and correct me, that in preterm you really have to have concurrent controls all the time because of the changes in care practices and how rapidly things can happen and because we simply don't know enough about those babies, for all of the practice environment of a preterm environment.

So if that's what you were saying, I would agree with that, that it isn't a one, two, three; it's really sort of a one.

DR. GARZA: I'll ask the group, as we conclude today, to please think about, B, because we pretty much addressed A. To a certain degree, we've addressed B, but I'd like to come back to that tomorrow after the public comment period, and then, depending on what we do with B, more explicitly, perhaps even returning to C, but maybe that won't be necessary, and then we'll pick up 5, and 6, and 7 at the remaining time.

I want to thank the group because, Dr. Sigman-Grant is absolutely correct, you guys have worked hard today, and you have one more chore before tomorrow

morning and that is thinking about Question 7. You may want to do that after dinner, rather than before.

I want to thank our presenters, because they've obviously made this discussion much, much easier than it otherwise may have been, and staff for getting this organized as well as it has been. So to each of the players, thank you, and we'll be back tomorrow at 8 o'clock--8:15. I've been corrected.

[Whereupon, at 5:51 p.m., the proceedings were adjourned, to reconvene at 8:15 a.m., the next day, Tuesday, November 19, 2002.]