obtain the longer-term information on delayed toxicity on quality of life and then extending, as I said earlier, into issues of survivorship. So, to me it's in that setting where the longer-term data collection is the most informative.

In fact, a drug that may be approved for use in the metastatic disease setting one could easily imagine might not be approved for use in the adjuvant setting because there might be long-term quality of life issues that counterbalance whatever benefit might occur in the adjuvant setting.

So, looking at this from the perspective of the advisory committee to the FDA, I think that that's how I see the issues playing out. It's not simply a matter of is it important to know clinically what the patient's quality of life is up to the moment of death. Sure. As a physician I'd want to know that. The real issue is how does that inform the regulatory decision making.

DR. CELLA: Thank you.

What I heard -- and then, Jeff, if you still want to say something.

To me then the question, based on what I heard, becomes why not. Why wouldn't we want the data? What I heard was that collecting the data does not require you to use it in an analysis, but it enables you to use it in

analyses that have an intent to treat approach where otherwise you would lose that opportunity from an analytic perspective. So, what I heard Diane and Nan say is you lose little by collecting the information and you gain opportunity to analyze the data in multiple ways that will help inform the potential value of therapy.

So, I'm trying to figure out, even from a regulatory perspective or clinical perspective, why you wouldn't want to have the data available for analysis as opposed to deciding up front that you're not interested.

DR. PAZDUR: One of the questions that I have is, are we confusing chronic toxicity with quality of life?

DR. CELLA: No, I don't think so because we're talking about --

DR. PAZDUR: So, I think what Jody was talking about was some aspects of chronic toxicity and evaluating the toxicity of an agent. Really, the evaluation of toxicity of any drug is not going to be answered only in one trial or two trials right when the drug is approved. It's going to be really a longitudinal basis even after the drug is marketed, et cetera.

The question that I have for you is, is the science of quality of life, as it exists in the year 2000, to a point where we can make regulatory decisions and that it is different that much in the analysis and the

distinction between quality of life and toxicity? But is quality of life in the year 2000, looking at this longitudinal basis, there to the point where we can start recalling drugs, making some assessment, and making regulatory decisions on that basis of just quality of life in the absence of toxicity? That's really kind of, I think, what you're asking here, David.

DR. CELLA: That was a really useful 3 o'clock direction or directive, but if I can take the risk of speaking for the committee, I wouldn't be here if I thought the answer was no. It would be a waste of time. I would say leave me to be in my office or laboratory or whatever you want to call it where I do my work to get it to the point where it's ready. But the truth also is it's only ready enough. It's not the best it can be. So, while we try to move that forward and make progress, there's a sufficient readiness as long as we're willing to accept some error that was commented on earlier off of Carol's point to accept that there's some error in the measurement, as there is in most of medical measurement. I can't think of too many things, other than survival where there's virtually no error, where there's not some error in the measurement. So, there's error in all measurement, but we've worked out enough of it so that the answer can be yes, and we need to sort out the guidelines for how to

proceed.

The other thing is that, unfortunately -- and this will probably always be the case -- patients and clinicians and quality of life measurement scientists cannot fully differentiate disease symptoms from treatment side effects. We can't do it. You can't do it, and people with cancer can't do it. Fatigue is caused by a disease and by treatment. Pain is caused by a disease, et cetera. So, that's always going to be a mix. We need to turn to the patients, we're saying, to get from them the sum of that mix and then make as reasoned a differentiation as we can between how much of this is disease symptom benefit versus toxicity offset or relative improvement.

If it's right on this point, Stacy, sure, go ahead, but there are some people in line, and you're second in line behind Jeff.

DR. SLOAN: I wanted to go back to a question which has kind of come around here. Why would you want to measure quality of life, for example, in advanced cancer patients after the treatment has failed?

And then you also asked, what is the science of quality of life measurement such that what do we know about what happens to quality of life shortly before death? Well, there are a few things that we do know. The most recent literature suggests, for example, it's intuitively,

one would think, obvious that all aspects of quality of life drop before you die, going monotonically towards zero. And that is not true, and there's empirical data to support that.

Physical functioning and health-related quality of life related to physical functioning are often confused with quality of life before death so to speak. Yes, those scores do drop, and if you look at studies that have been done by a number of folks around table, for example, in hospice patients, the physical aspects of quality of life, yes, will be going down monotonically and it's not surprising.

What is surprising, however, is some of the other dimensions, some of the other aspects of quality of life will actually retain their status from pretreatment and, in fact, get better for a lot of reasons, which I won't go into at this point.

But let me give you an example of where it was useful for us to assess certain aspects of quality of life in a cancer control study dealing with hot flash activity just recently looking at the use of SSRIs, selective serotonin reuptake inhibitors. There are some antidepressant effects here that potentially confound the issue of controlling hot flashes in post-mastectomy and post-orchiectomy patients.

It was important for us to know, even if a particular agent would have failed, would the impact on mood, for example, impact the patient's emotional quality of life, which we now know in the advanced days and later days actually does stabilize due, if for no other reason, to a response shift that happens in a patient accepting mortality and certain circumstances as long as they're relatively stable, not involving immediate hospitalization. Quality of life in many dimensions remains relatively high.

If some agent were to mess up that, especially again going back to the idea of how do we determine clinical significance -- if a priori we have a reasonable scientific hypothesis to suggest that, yes, there are some aspects of quality of life that may be impacted by this particular agent, even after it's failed, then that does offer us another question as to whether or not we should continue to give it even after it's failed or whether there will be lingering after effects.

So, again, I think there are situations where, yes, it would be useful, but again it goes back to what we said I think in the second session this morning. They have to be scientifically justifiable as opposed to just going on a fishing expedition, and I think that's what we're all kind of coming to hopefully.

DR. CELLA: Stacy?

DR. NERENSTONE: That was going to be sort of my point, that for instance you asked, David, well, maybe the rate of progression is less after you stop this drug, even though it doesn't have an effect. And I would say that we don't have the technology to figure that out. That would have to be designed at the very beginning of the trial looking at that specific question which I think is going to happen with the antiangiogenesis, those new agents.

I have no problem with collecting the data. I think you have to be very specific about how much data is really necessary because then it becomes an undue burden on both the physician and the patient. You do want to sort of know what you're going to do with it, even if it's just to assess for toxicity, which is what I think most people are saying. But to try sort of culling the data and then say, oh, see, look, we had delay of time to progression, without really knowing beforehand that that's what you were going to look for I think is really not valid. And just that one example, that's unknowable right now in terms of delay in time to progression.

DR. CELLA: I'd like to be clear and say that I was never suggesting this as a fishing expedition, nor was I suggesting it as a way of finding something that's there that we don't yet know about.

What I'm saying is I hear a stronger argument, frankly -- I sort of feel in the middle of this one because I am clinically trained and I do see patients. I feel a stronger argument from the statistics side that says you want to have the data because analysis options remain open when you have the data and the door is closed when you don't have the data. I don't see enough of a reason to not collect the information.

I'm not speaking to what you put in the protocol up front. If you put in the protocol up front that you're going to study time to symptomatic progression or that you're going to compare groups of patients up until the time they announce their progression, then that's what you study. You should stick with what you committed to when you start the trial, and I'm sure that will come through as a recommendation from us to ODAC and the FDA.

I'm just saying that I haven't heard enough of an argument for not collecting the information. Deciding about how to use it in an analysis is a different matter.

Any other comment?

We actually have really only talked about one issue, and clearly we need to, on that issue, come forward with some draft text and then see how it feels. I mean, one of the really rich things about this group is that it is admixture and a relatively small admixture of

statistically oriented people, measurement oriented people, and clinically oriented people. I think anything that we write and review is going to have to pass all three screens in order to be acceptable, and that's going to be an interesting process. But I think it will be fun.

We have other topics that were raised. We only have time for maybe --

DR. FAIRCLOUGH: Dave, are you going to the questions?

DR. CELLA: Well, we can do that too. Let's review the questions. That's a good idea.

DR. FAIRCLOUGH: I have two more slides.

DR. CELLA: Diane has two more slides that address the questions?

DR. FAIRCLOUGH: That address the questions.

DR. CELLA: Okay.

DR. FAIRCLOUGH: Well, to some extent.

DR. CELLA: The questions are what are the strengths and weaknesses of the major types of analytic approaches, i.e., longitudinal modeling versus univariate techniques.

And the second question is what analytic approaches should be taken to assess the type of missing data, informative versus non-informative, and handle them in the analysis. Example, various imputation techniques,

pattern mixture model, et cetera.

We've talked a bit more about the second, so why don't we start with the first? And no advertising.

(Laughter.)

DR. FAIRCLOUGH: The strengths of longitudinal and univariate, and I referred to it in some ways during the presentation. When we're talking about univariate in the context of repeated univariate t-tests or even repeated univariate changes from baseline, we're making those missing completely at random assumptions. And we also have the problem of multiple endpoints. So, I would strongly suggest that we don't use repeated univariate analyses in the presentations. I think there are easily accessible alternatives so that that should be something we eliminate.

When we go to longitudinal models and we use all the available data, we've relaxed our assumptions to the missing at random, and I think that should be the minimal criteria for an analysis.

There's another way to get to a univariate endpoint and that's to do some type of summary measure that takes the longitudinal data and converts it into a single measure. If you've got a very appropriate choice to that summary measure and you have a way to construct that summary measure in the presence of missing data, then that may be a very nice and interpretable strategy.

What's very difficult, though, is how you deal with the missing data and trying to set up rules to deal with it on a case-by-case basis. It will drive you absolutely nuts to do it, let alone to document how you've done it and justify it. So, you may actually end up developing a summary measure from your longitudinal multivariate model. So, instead of constructing an AUC for each individual, what you do is you get the quality of life versus time curve for the entire population, and that becomes your summary measure and it's a function of your estimate.

So, I think that addresses some of those issues.

Claire?

DR. GNECCO: Diane, for something like area under the curve, for instance, wouldn't you also like to see some kind of longitudinal modeling to put that in context? I always get the feeling that you could be losing information with an area under the curve approach.

DR. FAIRCLOUGH: You mean not seeing descriptively what's happening in the two arms?

DR. GNECCO: Yes.

DR. FAIRCLOUGH: Again, it really depends on the question, but I think that the descriptive plots of what those two arms look like in terms of the quality of

life over time may very well be appropriate to the interpretation.

DR. GNECCO: So, the more complete the graphical presentation, the better. Thanks.

DR. FAIRCLOUGH: I think the missing data is the one that we have been dealing with. The first part of the question asked about testing for it, and again reemphasizing that there's not a formal way of testing for non-ignorable data, but when we have missing data associated with morbidity and mortality, it's highly likely that we have -- and it just really requires that we really understand the disease process and the treatment that we're giving, and to do it naively, stick it into a package could get you in trouble. So, you need to think about what could be the things that you are measuring, make sure you're measuring them, and see if they're associated.

Then you need to do an analysis that's appropriate or at least a sensitivity analysis when it's suspected there's no simple and single approach. I think the really critical thing is -- and it takes about a day and a half for me to go through this in a formal way -- is to really understand the assumptions of each one of those methods.

Multiple imputation, if done very thoughtfully and with really good clinical information, can be really

powerful. If it's done in a naive manner and you happen to omit those important clinical correlates, it just makes the problem seem like it has gone away, but it hasn't. So, you really have a danger there. And we have to do sensitivity analysis.

So, the pattern mixture models -- I don't know how you implement those easily because you have to make such strong assumptions about what happens to the patients in that group, let's say, one of your patterns is those people who only had baseline. So, what do you infer about the quality of life of those patients after that? Those that drop out after the first assessment, what do you infer about their trajectories? There's a lot of assumptions there.

So, every method has a lot of assumptions, and if they're done well and justified well, then you have something meaningful that can inform you, but if they are done without understanding the assumptions, you just dug your hole deeper.

DR. CELLA: Stacy.

DR. NERENSTONE: Yes. I just have a question. There's always this discussion, when looking at the quality of life data, that data is missing. But in fact, we also know a lot about those patients because the investigators, to pull somebody off study, have to say why they dropped

out of study.

So, on the one hand, you're saying you have to infer about the trajectories they're taking. But in fact, you know certain things about them. Somebody has to decide whether they're progressing and so they're being pulled off because of disease progression, that there's toxicity from the treatment. That means that they're taken off because of toxicity, or both, or neither. They've moved away.

Is there a statistical way of folding that into the model? Because it's no longer the patient giving their quality of life, but it's also not missing data. We have information. We just don't know how to use that information in the mathematical models that have been constructed.

DR. CELLA: Well, I think it's guiding information. It's good information to have, and it can be used. Diane, if you want to explain details on how you might use that. But the short answer is that the more information you have, even if it's not patient self-reported status on a questionnaire, will help in the modeling.

DR. FAIRCLOUGH: It will help in the modeling, but you still have to make an assumption of how that information relates to the specific value of quality of life that you're going to infer that that patient has.

Does disease progression mean that they've dropped 15 points or does it mean they've dropped 25 points? And you're going to have to have some way to inform that. And if you have no quality of life measurements on the people that have progressed, then you can't make any inference, or if you have no quality of life on the patients who had a grade 3 toxicity, you're not going to be able to go anywhere with it.

DR. CELLA: Diane, I have a question, which is sort of a general question. Do you think that we're ready to -- this is kind of like along the lines of Rick's "are we ready" -- have some general guidelines? You're mentioning how there needs to be a sensitivity analysis. It's hard to you envision when there's missingness in the analysis which we can expect will always be there, at the very least because patients die, but also because they don't fill out the questionnaires.

Are we ready to have a list of options with some in a sort of preferred category? You have eliminated some, but among the ones you haven't eliminated, would you put them all in the same class? Would you say select this under circumstances A, B, and C versus select this one under circumstances C, B, and A, with sensitivity analyses done by the remainder or by two or three other approaches? I'm wondering if we're ready for a distilled kind of set of

suggestions, recommendations, if you will, for protocol development and planning.

DR. FAIRCLOUGH: The answer is no. I think one of the things is that there's so much development of methodology right now, and we're on such a strong curve. I mean, it's like the hot dissertation topic to deal with non-ignorable missing data. So, I would hate to write something specifically in guidelines that would, by the time we got it printed, there were three more articles that came out that would inform us. There might be better methods that then wouldn't be used because they weren't on the approved list.

The other thing that's probably even more compelling is I don't think that we can envision every scenario of every focused type of question and how it would play out in a particular disease or treatment well enough to make that kind of contingency plan. I think we can put guiding principles that will be informative, but I don't think we can have a checklist.

DR. CELLA: Maybe we'll start with -- I didn't mean to suggest checklist, but a list of sort of acceptable approaches under varying circumstances without necessarily putting them in rank order. But maybe we'll have these guiding principles and then be able to work from there to lay out some examples of typical phase III trial planning

situations, with this is a reasonable approach. And we may need to have several examples so that people don't get the idea that we're saying there's an imprimatur on this one approach. It seems to be part of the concern.

But it's going to happen. It is happening. There are trials being planned today where people want to know what's going to pass muster and what do I need to be thinking about to see to it that it's going to be acceptable given that there are many dissertations that still need to be done.

Donald, you had a point?

DR. PATRICK: I think I heard a lot more agreement about what isn't acceptable, and so that might be one way to go in terms of under what circumstances would this simply not be able to support the analysis, either the multiple testing issues or the missing data issues.

You also raised something at the end that I think has been a little bit of a problem all day long, rereading the questions in part one, and that is the phasing of these studies and what can be done prior to the pivotal phase III study to understand missing data issues and instrumentation and concepts and validation so that when the submissions come, you've got enough evidence so this isn't all hinging on one study. The pattern of evidence across multiple studies seems to me to be an

extremely important issue so that you've had more than one chance to deal with missing data on a particular therapeutic area and within a particular population that would support the approach when you put forward your key trial data. It's very difficult to specify in advance on a single study that this is going to work, depending on the power and the sample size.

DR. CELLA: One of the items for the last hour, this post-marketing and ancillary studies, kind of touches on that. What are the kinds of things that should be done, not necessarily within the pivotal trial but surrounding the trial before and afterward to be able to support a claim. And we'll come back to that. So, thank you.

Let's take a quick break.

(Recess.)

DR. CELLA: Well, we have had quite an interesting day. I think there has been a lot of stimulation around the table and a lot of interesting discussion. Several areas have been opened up. Very few have been resolved, but I'll remind you that at the beginning of this meeting I warned you that we had no intention of resolving anything today. And I think we've done a reasonable job of fulfilling that promise and spelling out the three primary issues.

But before we move to summarizing on those

three primary issues briefly, and then the other five issues and any others that are listed on the future plans, Dr. Beitz from the FDA will offer some reaction, perspective, and specific requests of the subcommittee. Julie.

DR. BEITZ: What I was going to start off with was just the idea that when industry comes to meet with us, we're often asked by them what instruments should they use in their studies or, as a corollary to that, is instrument A or B or C acceptable to the agency. Very typical question. I would just like to comment on some of the things I heard earlier today that may speak to this point.

Regarding the overlap of symptoms and health-related quality of life, we heard this morning that symptom measurement is a good place to start but may not tell us all that we would like to know. So, perhaps a topic for another day, a further discussion, what are the clinical settings that are appropriate for a symptom-only study versus what settings would be appropriate to have additional information with health-related quality of life instruments.

As an example, it could be that perhaps where you really need the additional information from health-related quality of life might be in an early submission, your initial pivotal studies where you don't know very much

about the drug. However, once the drug is approved, on the market, and you have prior information on its toxicity profile and you're submitting additional studies for supplemental indications, then perhaps those studies may only need to have certain symptoms evaluated. That's just a thought that I had that you could toss around.

Another thing that I heard that I found fascinating was Dr. Sloan's discussion of how when he was working on a colorectal cancer trial, that he got together with the clinical trialists to review the questions that were most appropriate to ask colon cancer patients, and from there, going to the instrument that was of interest, the FACT in that case, and to see how much of these questions were actually covered by the instrument. And I guess most of them were, except for a couple of items which were then written into the protocol.

I found this very interesting, and I just throw out the question, how frequently is this kind of comprehensive exercise actually undertaken? And if it were done routinely and comprehensively, would we actually see different instruments proposed to us in trials? And would we get potentially different results?

That brings me to the idea of results, and I would just echo Diane Fairclough's question to us today which was, now that we have the data or not, what do we do?

I'm not a statistician, but what I did hear was there is no single approach that will probably apply to all the situations that arise in analysis, but that perhaps what could help us are delineation of the pros and cons to the different approaches. I think that if we understood what the strengths and weaknesses were of the different approaches, then we could assess these approaches when they're done in submissions that are given to us and also we could critique our own analyses of these data.

So, I would just like to conclude that what I learned today was that to get a handle on the value-added health-related quality of life, if this added value exists, it would be important to have a continued dialogue between quality of life researchers and clinical trialists, and I believe that this meeting is a good start toward accomplishing this.

DR. CELLA: Any questions or comments?

(No response.)

DR. CELLA: Thanks, Julie. It was helpful. Carol.

DR. MOINPOUR: I've been talking to several people over the course of the day and there is another committee within the FDA that's looking at quality of life assessment and it would seem to make sense for the two groups to have some way of dialogue or communication. That

might be useful.

DR. CELLA: You're referring to Lori Burke and that activity?

DR. MOINPOUR: Yes.

DR. CELLA: Well, Lori was involved in the planning of this meeting and I think is not here because of a conflict, but my assumption is that she'll be at the next meeting unless there's another conflict. Is that correct, Julie?

DR. BEITZ: Right. I think there are folks actually in attendance here, including myself, who sit on this other committee. There is a committee looking at trying to write guidance on this issue, and it is our hope that at some point we can actually present the document that we've got to this group for comment down the road.

DR. CELLA: So, you'd like to wait to present that to us because you want to get it into a condition that you're comfortable with internally rather than having us work on it with you, which is fine. I just want to clarify why. Because the other way to do it, of course, would be to show it to us sooner for comment rather than later.

DR. BEITZ: Right. I don't want to speak for Lori. I think she's ready pretty soon.

DR. CELLA: Okay.

Along those lines, there are other documents

that we will circulate to all the subcommittee members. One of them is from the International Society for Pharmacoeconomics and Outcomes Research, and it's their position paper on health-related quality of life claims, which is also on their web site. I've just downloaded it from their web site, and we'll see to it that it's copied and distributed.

The International Society for Quality of Life Research is not yet circulating its draft document that is intended to advise on claims in labeling, but when it does, I'm sure we'll be interested in seeing that as well.

And PhRMA also has a document. I don't know. I think it was due in the summer, but I haven't seen it, and I don't know if it's available yet. Maybe others do. Does anyone know if the PhRMA document has been disseminated? Anyone in the audience know? Yes, please. There's a microphone there. If you could speak into that, I'd appreciate it. Could you tell us who you are, please?

MR. WILLKE: A long way for a brief comment.

I'm Dick Willke. I'm with Pharmacia & Upjohn and also on the Health Outcomes Committee of PhRMA.

I'd say we're just about on our final draft of some commentaries from a meeting that was held in March of 1999 that was sponsored by our group, but we worked with the FDA DDMAC group to organize it. There are some results

on many of these same issues that are in this working draft. I can't give you a day when it will be available exactly, but it should be very soon.

DR. CELLA: Would it be more appropriate for him to correspond with you or with me, Karen, just to inform us about when this is a final draft that can be circulated? Would you be willing to share it when it's available for circulation?

MR. WILLKE: Certainly.

DR. CELLA: Could you share it with me and then I can ask that it be distributed?

MR. WILLKE: Be glad to.

DR. CELLA: Thank you.

I believe that there's no use in our operating in a vacuum, so I think it's better to have these things than to not have these things. This is just a downloaded text from the ISPOR web site, and then there's this paper that Carol Moinpour actually mentioned this morning by Nancy Leidy and Dennis Revicky and Bernard Genest, published in Value in Health last year, on recommendations for evaluating validity of quality of life claims for labeling and promotion. So, we'll circulate that as well. So, look for a package to come in the mail for you to review between now and the next meeting.

We have another half hour to discuss future

plans. The first thing I'd like to do is just very briefly review the fact that we've covered these three major areas: definitional issues, clinical significance/interpretation, and data analysis. I have volunteered myself to take a stab at a first draft of a brief summary targeted toward guidelines, tone if you will, that I'll first circulate to the presenters and discussants for their comment and then after getting those comments back and making modifications that I'm sure will happen, then circulating to the entire subcommittee. I'll do that well in advance of the next meeting in June. We'll do it through Dr. Somers so that she can circulate it to anyone in the agency that is involved and interested. So, that's just a review of what will happen between now and the next meeting by way of clarifying and optimally summarizing what just occurred today.

The other items I wanted to touch on, as we consider where we're going to go with the next meeting, which by the way will, we believe, either be June 5th or June 7th. My hope was that that would be determined by the end of today, but I guess that was not realistic. Karen would like to know. It's dependent upon ODAC which will be either the 6th and 7th or the 6th only. It's not clear how many days it will be, or the 5th and 6th or 6th and 7th. So, it's dependent upon that.

But to the extent we inform them about our conflicts, it might also actually be dependent upon our schedules as well. But that all needs to be coordinated by Karen. So, please let Karen know if you have a conflict on the 5th or the 7th, which are the two most likely next dates. If you don't have a conflict, let her know that too. Thank you.

So, as far as substantive future plans, there are these five areas that represent some things that we knew we wouldn't have time to get to today, but are also very important. There may be others to add to the list, and I encourage you to do that.

The first is when should one study health-related quality of life. We've touched on that through the course of discussion. Julie asked the question about when would measuring symptoms be enough, and implicit in that is when should one study health-related quality of life. So, we need to have some discussion about the parameters that should be guiding our thinking and the FDA's thinking about when it's appropriate to be looking at any of these endpoints.

Second is all of the many issues that come up with study design. The question about timing was asked, which is a design question. The question about whether you assess quality of life after an event as in progression was

also asked. These are design issues. There are several other design issues that were touched upon, and clearly we need to delve into more specifics on those.

Would anyone like to nominate any design issues that you think are pivotal, if you'll excuse the term?

DR. GNECCO: What comes up very often is when you do come up with an assessment schedule, if it isn't very closely adhered to in the analysis, you can have a real problem because of measurement error. So, that has to do with the timing. If they're too widely spaced in time, you have a greater likelihood of people being outside of a large window. So, maybe thinking about time windows associated with the measurement schedule.

DR. CELLA: So, it's both the timing, the frequency, and also the acceptable margins around which you can receive data and analyze it.

DR. GNECCO: Exactly.

DR. CELLA: Well, it seems to me that timing as a topic area is broad enough topic area to merit an actual set-aside discussion. I see some heads nodding around the table. Just as we had three main topic areas today, why don't we assume that the next meeting's agenda will include a specific discussion of timing?

Questionnaire validity and acceptable modifications. This is an area that really apparently also

comes up a lot. Julie started with this, saying that this is one of the most frequently asked questions that they get from sponsors. Embedded in there are things like how do you know if a questionnaire is valid. I hope we don't need this to become a sort of basic primer or 101 course on validity and reliability. I think the FDA has just had a seminar on that, and there are certainly available texts that we can refer people to. But I'm concerned if we get into too much on that, it will divert us from our charge really.

But nevertheless, it remains important to push the envelope, if you will, on determining what are the minimum acceptable requirements for a proposed set of questions. Does there need to be a published paper on a scale that's being proposed? When is it acceptable to extract items which came up in the discussion today, and if one did that, what are the sort of acceptable ways of doing that? What are the things to be concerned about? Is this an area that around the table we think is something we're ready to tackle in June? I think it's something we're going to have to tackle.

Carol thinks it's easier. I think it might be harder.

(Laughter.)

DR. CELLA: So, validity and acceptable

modifications. Is there a better term anyone would like to nominate for that?

DR. SLOAN: How about non-standard use of standardized questionnaires?

DR. CELLA: Okay, or new uses? Old wine in new bottles?

(Laughter.)

DR. CELLA: Just acceptable measures, she said. Minimum requirements for acceptability of questionnaires.

VOICE: Psychometric integrity.

DR. CELLA: Okay, psychometric integrity.

So, that's a probable second choice. I just want to run through them before we commit.

Post-marketing and ancillary studies. Now, these are not necessarily the same thing, and maybe they don't belong on the same line even.

Let's start with ancillary studies actually because this has come up today a couple of times. The idea here is not everything has to be attached onto a pivotal trial. Arguably it should not necessarily be attached onto a pivotal trial, but may provide important supplementary information. It might tie in with this other area of psychometric characteristics, such as new validity data on a questionnaire or the value or preferences that people place upon certain symptoms, whether they're symptoms of

disease or side effects of treatment, or the extent to which symptoms drive functional problems or changes. These sorts of things can be studied outside of the trial for information that can be given in support of the application.

It seems to me this, among the list, might not be necessarily ready for June but we may need to cover that.

Rich?

DR. SCHILSKY: I was just going to say to me more important than post-marketing studies might be to have some discussion of what kinds of information could potentially be collected, say, as part of phase II trials that would be useful in facilitating hypothesis generation for testing in phase III trials going forward. Since we've been stressing the importance of hypothesis-driven research in this area, it seems to me that you have to be able to have a minimum amount of information to enable the development of a hypothesis. So, what could you do during phase II that would facilitate that?

DR. CELLA: Yes, I agree. I mentioned it earlier and I think that these kinds of ancillary studies that I referred to really don't belong with the post-marketing line. So, I just created a new line, almost like pre-submission ancillary studies, because so far we're

talking about the kinds of studies that might accompany a submission rather than anything that would be post-marketing. And that would include the phase II preparation studies.

Did I get that right?

DR. SCHILSKY: Yes. I guess I'm not sure that they should all be lumped together. It seems to me they address somewhat different things, but they're all important. I'm just thinking in terms of trying to ultimately provide some guidance to industry and investigators in terms of how to develop these studies so that at the end of the day, there's a cohesive and coherent data set sort of in parallel to what you would get in your clinical data set. There's phase II data and there's data from randomized clinical trials.

DR. CELLA: Let's come back to that as we plan for the next meeting.

Reporting and labeling. How about this area?

DR. PAZDUR: I think that's important to some people in the FDA, but I think of the topics that we've described, I think we're going, not in a chronological order, but order of importance to us as far as the division goes.

DR. CELLA: I'm not sure the mike picked you up there. Did you get that? Would you like me to repeat it?

He said thinks it's important to some people in the room, but not of as much direct importance to ODAC and to this division of FDA.

I think if we could keep that as a future possibility because some of these things we do are going to logically flow right into labeling implications and we don't necessarily need to specifically deliberate on labeling. I think it would be an automatic flow from some of the work we do.

Carol, did you have a comment?

DR. MOINPOUR: Not about that.

DR. CELLA: So, that's a later.

Are there any other areas that we'd like to hear nominations for? Stacy.

DR. NERENSTONE: In trial design, you didn't say anything about randomization. Is that because implicitly they're all going to be randomized or that they're not going to be randomized, or you don't think that's an important concept for us to talk about? Or should it be added to the list?

DR. CELLA: I think it probably should be added to list. I had the assumption, apparently erroneous, that pivotal trials that you look at are all randomized trials. So, I was assuming something incorrect. So, it should be on the list.

DR. DICKERSIN: Rich mentioned the phase --

DR. CELLA: The phase IIs, yes. But I thought he was mentioning it in the context of work that could be done in preparing --

DR. SCHILSKY: There are also occasions when we're asked to consider things that are being proposed for so-called accelerated approval which are oftentimes non-randomized trials. Of course, there we pretty well are focused on a surrogate for clinical benefit, namely response rate. But I think even there, having information about the quality of life of the patients related to whether they're responding or not is important because there's a regulatory decision to be made about whether the drug should be marketed based upon this accelerated approval guideline, and that doesn't require randomized trials.

DR. PAZDUR: Yes. It doesn't require randomized trials, but it does require a commitment subsequently to demonstrate clinical benefit after the drug is approved. So, that application would come later probably when we make a phase IV commitment and should be an integral part of that phase IV commitment after the accelerated approval. So, it still would play a part, but probably a step later in an accelerated approval. You're just kind of changing the time line here.

DR. CELLA: It seems logically then, assuming that one can produce a reliable indicator of change in an individual and then that can be used to define response, that quality of life response can be taken to be evidence for rapid approval in a phase II setting and even clinical benefit if it was then looked at in the randomized trial. This is a conceivable scenario. It hasn't happened.

DR. PAZDUR: In general, what has happened in these situations is taking a look at a response rate, for example, as a predictor for approval. What you're saying in an accelerated approval is that you're approving the drugs on a surrogate endpoint. If you've already demonstrated clinical benefit in your mind, then that leads to full approval of the drug.

DR. CELLA: Yes, I'm just thinking. There may be a need for a minimal paradigm shift there in that if this is bought, that a quality of life benefit is meaningful and sufficient for approval, that one could conceive of a situation where you're actually getting phase II data that support rapid approval that actually, by the letter of the law, would also support full approval, although you'd then be approving something on a phase II study. So, there may be a little bit too much unprecedentedness to that, but it seems to me that it's an area that is worth discussing as long as it doesn't get too

philosophical.

What do other people think?

DR. SCHILSKY: I don't think we should get hung up in this committee on the sort of the letter of the regulations. That's the FDA's problem. I think what we should think about is the extent to which we can get informative data in a phase II trial that would help in the evaluation of an application.

DR. CELLA: Well, having heard that, let me suggest three new areas, not new areas, but among what we've discussed, three areas that we can draw out for the agenda next meeting. And we don't have to have three. We could have two because we'll be revisiting these other three. So, we don't need to pile on ourselves.

Sorry. I think I missed some people that want to say something. Diane and Jeff.

DR. FAIRCLOUGH: One of the things that I think would help the committee is if we had a better idea -- we didn't get any background information on the types of drugs that are reviewed by this committee. I don't know whether supportive therapy would be under the review of this committee or whether it's just limited to drugs that would have either a tumoristatic or tumor destruction thing. So, I don't know whether something like the growth factors would ever be -- and so, I think that would help us focus

when we're trying to answer the questions. I think it would help us to know a little bit more about the process and the group of drugs that we're working with.

DR. PAZDUR: I think the next time we could bring that up because it's quite complicated because some drugs are handled by CBER, biologicals that could be considered oncology products as such but not drug products, supportive care medications. For example, antiemetics would not be handled by this division. Pain medications would not.

DR. FAIRCLOUGH: It would help just to have a short summary of that for us to help define things a little bit.

DR. CELLA: We were chatting -- I'm sorry -- side-barring a little bit, but I think you were asking for a summary of the kinds of submissions that have been received over the recent past, and Rick said that that can be provided, so something that would be hopefully not the full transcript of the meetings but a distilled summary of the meetings. The full transcript is on the web, I'm told. We can tie it together with this packet that we're going to send around.

DR. SLOAN: I wonder, for example, you said that pain meds and antiemetics would not be included.

DR. PAZDUR: It's done by other divisions.

DR. SLOAN: What the other divisions have by way of guidelines as well might be helpful to make sure that we're not saying things totally at odds, for example.

DR. BEITZ: There is no formal FDA guidance on this issue.

DR. CELLA: We're blazing the trail.

DR. BEITZ: You're making it up as you go along here.

DR. CELLA: We'll see how much this can help their work.

DR. PAZDUR: But I think I would not want to lose focus by going in multiple directions here of multiple drugs. I think specifically what we're interested in is oncological drug products, and let's probably say for the sake of it, drugs that are producing either an antitumor activity or a cytostatic activity because I think since we do not handle pain medications, antiemetics, that probably is more of an issue for the agency in toto and this other group that is looking at quality of life. I think if we get diverted in multiple different areas of how to develop pain medications, how to develop --

DR. CELLA: I think the interest was just to see what's available to get some ideas and perspective. But if there's nothing available, then that makes it easy, at least on one level.

Dr. Justice, did you want to say something?

DR. JUSTICE: Yes. I just wanted to go back briefly to the question of blinding of trials, and this is a design issue. The question might be how do you interpret unblinded trials where quality of life is an endpoint. When can you interpret them? We deal with this day in and day out. We're told we can't blind the trials. So, I think we have to come to grips with two scenarios. One would be a double-blind trial and the other would be our typical oncology unblinded trial. What are the different considerations?

DR. FAIRCLOUGH: Wouldn't some of those considerations be similar to unblinded evaluation of toxicity, looking for certain symptoms?

DR. JUSTICE: Well, I mean, it depends. The typical toxicity is assessed by the physician, but if you're talking about health-related quality of life assessment of toxicity by the patient, sure, I think the same -- things like asthenia or fatigue, as we said before could be disease toxicity.

DR. FAIRCLOUGH: Yes. I mean, if the physician doesn't ask, he doesn't know. There's an assessment bias that could be conditional on the physician's knowing the particular drug, and so it's the same issue in some ways.

DR. PAZDUR: Correct.

DR. CELLA: If I suggested that we next time focus in, not necessarily completely at the exclusion of other items, but focus in on timing of assessment, acceptability or minimum standards of questionnaires, and kind of how to develop a pre-submission package that tells a coherent story, sort of along the lines of what Rich was saying, would those seem like a reasonable three? It sets aside for the moment issues like blinding, reporting and labeling, anything post-marketing, and any other ancillary studies or study design issues that aren't covered by those.

Does it seem like a reasonable course to pursue? I don't think we can even imagine adding more than three items, and that may be too much. We may have to trim that because we still have to review the first three.

How about only two? Timing and acceptability and the prestudy stuff hopefully can flow right from that along with issues like blinding. Sound good? Okay.

Creating a pre-submission story.

Yes, Kay.

DR. DICKERSIN: When you talk about blinding, it's not quite related but it's kind of related. At lunch some people were mentioning to me, who have been on trials, about the patient being so afraid of being kicked out of a trial, that they don't give the true story about their

quality of life experience or that they make things sound better than they are. It kind of fits in with validity. It's not a blinding issue because even if they're on placebo, they're going to say everything is great.

I don't know where that fits in, but it fits in somewhere in what you're talking about. And I don't know how to get at it, but maybe part of assessing quality of life, if we mean it to be meaningful, we have to reassure the patient that it will not affect their participation in the study and that that would need to be part of what we're talking about too.

DR. CELLA: In measurement terms, you're referring to a concern about response characteristics, things that influence the way in --

DR. DICKERSIN: Systematic misclassification problems.

DR. MOINPOUR: Social desirability and affecting the response that's related to that.

DR. CELLA: And hope.

DR. SLOAN: I don't know if, along those lines, you want to talk about providing QOL data as part of the eligibility. I know that has been an issue that the cooperative groups have dealt with, and I'm not sure we've got a final answer yet per se. We've got guidelines. But that kind of comes around that as well.

DR. CELLA: Well, we need to close up now. I want to thank everyone for coming very much. We were all engaged in interesting and active discussion. There's a lot of work to be done. I think we set a good course. You'll be hearing from us in the mail and on the Internet to clarify the summary of this meeting and the agenda for the next meeting very soon. Thanks.

(Whereupon, at 3:55 p.m., the subcommittee was adjourned.)

- C -

- E -

- F -

- G -

- H -

- J -

- K -

- N -

- O -

- P -

- Q -

- R -

- T -

- U -

- V -

- W -

- X -

- Y -

- Z -