CHAIRMAN GLASSMAN: Dr. D'Orsi?

DR. D'ORSI: I just want to ask a question to the gentleman from, I guess, Siemens.

How did you arrive at two false positives, which seems unbelievable with the territory you cover, and were there any false negatives, and how did you validate each?

DR. GUPTA: Absolutely. It is unbelievable in some sense that a study that comprises hundreds of individual image slices -- The technology has matured to the point that you have two FDA approved products in the U.S. today, where lung CT CAD performs at that level.

DR. D'ORSI: What was your false negatives with that two false positives? How many, just an idea?

DR. GUPTA: The lung CT study is carried out a little bit differently. I think you have a synopsis of the study in your material as well, and Dr. Naidich referred to that MRMC study. So it is actually study false positive rate and the positives.

CHAIRMAN GLASSMAN: Any other questions for our speakers? No? Then we will thank you all very much.

We will now move on to the general

discussion portion of the Panel's deliberations, general discussion of lung CAD.

I think if I may begin just to say that -- to reiterate what I said briefly before, number one, a lung biopsy is not a breast biopsy. It is a much more serious procedure. So as has been suggested by the Panel, false positives are really a critical issue here.

The other critical issue that I see is the satisfaction of search issue that is evident in some of the data for both chest X-ray and CT, and that if we miss a clinically relevant disease because a CAD mark distracted us, that that is a problem that may have to do with the concurrent reader versus second reader, and the issue for the agency of significant risk that I think, at least in my mind, is there.

Are there other comments about my comment or general comments? Dr. Steier?

DR. STEIER: Yes, comments from a pulmonary clinician or viewpoint. I was making an inventory of what I think are the positive and the negative features of CAD for CT scan and chest X-ray.

I think, as has been pointed out, since the

jury is still out on screening for lung cancer in asymptomatic individuals, it is different than mammography and colon and some of the other things we've talked about. I don't know if it is quite the Hatfields and the McCoys, but maybe so.

The number of false positives is concerning, whether it's 10 per CT scan or two per CT scan or somewhere in the middle but in clinical practice, that always leads to a great deal of cirrus, as they say, trying to track down all kinds of spots on chest X-ray or a CT scan, most of which will turn out to be nothing, but can go as far as lobectomy, you know, biopsy, open lung procedures, et cetera.

So definitely there is an invasive component that has to be considered, and it sounds to me like the more CADs we do, the more of these types of nodules we are going to find and have to deal with.

Also, the effect on the reader has been mentioned, and once we start highlighting certain areas, it does distract us away from other areas, and things could be mixed.

I think there are some positives, though we

will definitely pick up more nodules and in the unusual case where there is a significant more than 8 millimeter nodule that does have to be dealt with and may turn out to be early lung cancer, which may be in that very small percent of the population where that is curable, that would be helpful.

It is also interesting that there is some data to support both concurrent and second reader algorithms which is helpful to know as well in terms of the proper labeling.

So those would be some of the positive and negatives that I see.

CHAIRMAN GLASSMAN: Other comments about lung CAD in general before we move on to the questions?

Well, let's move on then to the questions, and I am sure other issues will come up as we go through.

We are going to begin our focused discussion of the FDA questions. Copies of the questions are in the meeting handout and on the tables outside the conference room.

We are going to start with question L1. Can we make that any bigger, Sunder, so everybody can see it?

There we go. Actually, it was so I can see it.

Establishing ground truth, whether disease is present and, if so, its location and extent, is crucial for the evaluation of performance of a CAD device. Please provide your recommendations for defining ground truth for lung CAD devices.

Doctors Carrino and Garra, do you want to start off this discussion? Which one of you wants to go first? Dr. Garra?

DR. GARRA: First, I was thinking, oh, well, you might have to have biopsies, but I don't know. We are just looking for detection of lung nodules here, and we are not trying to necessarily characterize the nodules as to what pathology they have.

So I don't want to trivialize this, but I think that we could establish ground truth by just using an imaging modality and looking very carefully at that area. I'm pretty comfortable with that level.

So I would say, for instance, if it was a chest X-ray and it had flagged a nodule, you could establish the ground truth by using a CT, for instance. In the case of a CT device, I think a consensus panel

with a high resolution CT would be sufficient, for instance. But I would be interested in hearing other comments on that.

DR. CARRINO: I agree with those, and for CT you could also extend it to a follow-up examination.

DR. GARRA: Yes, follow-up is also useful.

DR. CARRINO: So expert panel and follow-up examination.

CHAIRMAN GLASSMAN: Would you or any other Panel members include for the expert panel the knowledge of the findings from the CAD when they review the CT? Is it a blinded CT reading or a reading where the marks are known?

DR. GARRA: I personally would say that a panel has the information of what the CAD found, and they are looking specifically at that area in great detail. Specifically, maybe like with the modern scanners, you have the 0.9 millimeter thick slices, and you look at that original data and make very thin slices to really go through the area in detail.

DR. CARRINO: Yes, marks are known.

CHAIRMAN GLASSMAN: Any other comments about

the definition of ground truth? Yes, Dr. Tourassi?

DR. TOURASSI: I think we both have the same comment, but there are enough studies out there on CAD development where a panel consensus or follow-up CT studies were more than sufficient. I don't think we need to put the CAD as part of the truth in. But I don't think we should put the CAD marks in the loop as part of the truthing.

Actually, that was the case with a LIDC study that Dr. Clark mentioned before that was done from NCI where expert panel -- painful, but it was more than sufficient for that.

CHAIRMAN GLASSMAN: What do other members think? We've now got a two to one split on whether the experts should know the results of the CT scan or not.

DR. SAHINER: I think LIDC established some methods for providing ground truth, and in their method what the radiologist first said was to do an unblinded -- sorry, a blinded read where they were blinded to the readings of the other radiologist. So they marked each of the nodules that they found, and then in a second unblinded read, then the findings of other radiologists

were shown to them, and then they would either agree or disagree with that.

So in that unblinded mode, I think you could include the CAD marks and see if the radiologists agree or not, because there is -- also, I want to mention that there is a large variability between radiologists, even when they are reading in the unblinded mode when they look at other radiologists' findings, whether they define what some other radiologists found as a nodule or something else.

So there is a huge variability. I think it is somewhat problematic then if this paradigm is used, blinded and then an unblinded read. At the end, how do you merge the readings from multiple radiologists?

So another option could be to then have them sit together and do a consensus read, but again, I see that it is not very easy to do because the number of radiologists needed to do such a truthing also needs to be, I think, more than two.

CHAIRMAN GLASSMAN: Dr. Dodd and then Dr. D'Orsi.

DR. DODD: I just wanted to point out,

though, that if you include the CAD marks in the truth definition, you are going to bias yourself in favor of the CAD markings. So in general, from a pure statistical standpoint, the gold standard or ground truth should be totally independent of the test under study.

I understand there are practicalities that need to be addressed here, but I would be in support of Dr. Tourassi's comment.

CHAIRMAN GLASSMAN: Dr. D'Orsi?

DR. D'ORSI: Well, if the task is detection, I can understand that, but the task is validation. I think, with that task, to prove whether there is something there or not, I think, you reduce bias of a reader by directing him and just asking the question, is this a real mass or not. It's not a whole detection process again.

I think that if you do have a detection problem, you have the possibility of bias. They may see something other than what is marked and said oh, my stuff is really true, that's garbage.

So I think, if you focus them on what the task is, I think you would get a better result.

CHAIRMAN GLASSMAN: But isn't -- oh, Ms. Brogdon?

MS. BROGDON: Thank you. I just wanted to point out that the Panel seems to be assuming that you are talking only about a computer assisted detection device, and you may also need to include computer assisted diagnostic device in here where the endpoint may be -- or the claim may be malignancy rather than a nodule. So you might want to broaden your discussion.

CHAIRMAN GLASSMAN: Thank you. We will do that right now. One comment, though, about Dr. D'Orsi's comment is, really, we are dealing here with both detection and validation.

If the CAD devices misses lesions, that is just as important as finding them for here. So I think it is not just validation. Dr. Rosenberg?

DR. ROSENBERG: I was wondering whether validation would also include size criteria, because clearly, nodules smaller than certain thresholds are not important, and how does that validate it in terms of the size measurement?

DR. MITTAL: One of the issues about

validation and using a grounding truth for nodules for CT scan is probably a PET scan which is a physiological, obviously, examination.

It will be interesting to have Panel's opinion in patients that are found to have nodules on the CT scan and use PET scan as a validation.

DR. STEIER: I think you can use several different modalities to look at how well the technology is working. Occasionally, we have biopsies which, where possible, is the most meaningful information, of course, PET scans, which can be used to sort out, in most cases, malignancy from benign, as well as the panels and serial CT scans, et cetera.

So I think we can use whatever tools we have available to evaluate the new technology.

CHAIRMAN GLASSMAN: So let me just make one comment, and then we will come back to Dr. Tourassi.

So we've got two issues here that we are discussing. One is detection, and the other is diagnosis. For diagnosis, which in many ways may be the simpler of the two, ground truth, to me, is either pathology, PET scan, or follow-up scan.

I don't think we can expect pathology on every lesion although if the computer aided device suggests a malignancy, I think most of these will either go to PET scan or to surgery or lung biopsy. So we will get that.

Let's, if we can -- Dr. Tourassi, you had a comment, and then if we could just focus for a few minutes on the diagnostic end and sort of nail that down maybe, and then move back to detection.

DR. TOURASSI: Yes, I was actually going to summarize that ground truth could be expert panel decision based on whatever information is available for every case. So is it a follow-up CT? Is it a PET study or is just the image they have in front of them, but there's three people that need to decide and agree on. That would be the ground truth.

CHAIRMAN GLASSMAN: Dr. Garra?

DR. GARRA: I just wanted to -- yes, if we have broadened it to include diagnosis, then I agree with you exactly, that you are talking basically about a biopsy or a follow-up that shows no change in the case of the most benign nodules or perhaps a PET CT to look for

activity, and then probably followed by a biopsy.

So it is quite a different question from just detection of nodules.

CHAIRMAN GLASSMAN: Any other -- yes, Dr. Carrino?

DR. CARRINO: I would just agree with Dr. Tourassi's assessment. The expert panel takes into account all this information and then decides, and that is used for other diseases that may not have a pathological endpoint like multiple sclerosis and those things. So I think that is probably the best way.

CHAIRMAN GLASSMAN: Another issue with lung CAD that I have identified is the issue of follow-up studies. There are two things there. One is nodule growth, and the other is nodule morphology.

Morphology is used in the initial study when you are worried about carcinoma. If you are trying to do diagnosis, but it becomes even more critical in the growth issue because nodules that grow and are smooth and oval or round are one thing, and nodules that grow and are irregular are something else entirely.

I think I would be interested in the Panel's

thoughts on how much testing that portion -- if that is included in the CAD device rather than in the non-CAD portion, the CT device -- as to whether that needs to be validated as part of the application and, if so, how do we ask that that be done?

Anyone? Dr. D'Orsi?

DR. D'ORSI: You are asking are there specific morphologic criteria that the CAD can identify on the diagnostic end to give you an accounting of whether you think it is malignant or benign? Is that the sense?

CHAIRMAN GLASSMAN: Well, that is one thing. The other is the accuracy of the volumetric comparisons that become critical for the paradigm of growth. It is growth and morphology on follow-up.

We need to make sure that these systems, if in fact it is in the CAD piece -- if it is not in the CAD piece, it is not an issue. Like we'll talk later a little bit probably about breast MRI where a lot of that is functional and measurement. It is really a giant measurement package is what it is.

If it is in the CAD piece here, I think that

we need to be sure in some manner that it works as intended.

DR. D'ORSI: Yes, I agree with you. If it is going to be extended -- I thought this was purely a detection task, but if you are going to add diagnosis, I agree 100 percent that you have to be very critical, that it can measure size difference accurately, that it can measure morphology accurately, et cetera.

So this is a totally new thinking for me.

CHAIRMAN GLASSMAN: Dr. Garra?

DR. GARRA: At this point, yes. I mean this stuff is just coming on the horizon, these nodule growth packages that you can see, and they give you doubling times and things like that. If it is included in the CAD package, it does need to be validated.

I know of no good way for a panel to do that. I don't think humans are able to actually do that. That's why these are useful, and you are going to have to have phantom experiments with various shapes and a fairly complex set of phantom experiments that show that it is valid in a physical sense, that it is actually measuring the edges of lesions of various types and

shapes, and able to categorize types and shapes.

We have had to use phantoms before a lot, especially in ultrasound, and they can be designed. There is a pretty sophisticated industry to do that.

DR. STEIER: From a clinical viewpoint, this is going to be counted on as reliable information so if it is going to be provided and in the package, it has to be validated. I agree.

CHAIRMAN GLASSMAN: Any other comments about ground truth? Yes, Dr. Bourland?

DR. BOURLAND: I agree that these quantitative assessment of images, whether sequential imaging, follow-up images, et cetera, that that would need to be validated. Phantoms are very powerful for this. They tend to provide fundamental information on how the algorithms in devices operate. So in some sense, they stand a little bit more in the standalone evaluation of the system. There are limitations in phantoms, but they can be made static. They can be made dynamic, et cetera.

Things such a whether you include the boundaries or exclude the voxel size, et cetera, are all

aspects to consider.

CHAIRMAN GLASSMAN: Yes, Dr. Abbey?

DR. ABBEY: I just wanted to say, I heard something in the idea of consensus panels that made me a little nervous which is that there is high reader variability. So in the absence of additional information and a follow-in scan, a PET scan or a biopsy, I am actually a little concerned about a consensus panel that only has the image there in front of them to use.

If readers are really that variable, does a consensus panel really get at ground truth or do you just test for consensus?

DR. CARRINO: That depends. So if the purpose of the consensus panel is to say, you know, for detection is this a nodule, yes or no, I think that is a valid way to do it.

If you are looking at other diseases like cancer, then these expert panels take in a bunch of different information, like usually these patients have either longitudinal follow-up or they have some other -- There's different levels of truth. There's gold standard. Is it silver standard, bronze?

So you have to put all that together to determine what the ground truth is, and I think that is the best way to do it. That is the most reasonable way that we have today to establish a ground truth.

CHAIRMAN GLASSMAN: Dr. Dodd?

DR. DODD: You could consider incorporating the variability of the consensus panel into the truth evaluation in some sense.

I also think, you know, consideration needs to be given to the way truth is defined, even with the consensus panels at majority vote, et cetera. You know, there is a study many years ago that shows that, you know, I think there were four different ways of defining truth, and each one gave different estimates.

So if nothing else, when you are comparing things, we need to agree that truth is defined in a consistent way.

CHAIRMAN GLASSMAN: Does anyone want to define truth for a consensus panel of three or five? No takers? I don't have an answer to that one either.

Let me ask another question to the Panel, and that is what about non-nodule findings? Should they

be included in ground truth because one of the potential issues is the satisfaction of search and the missing of clinically relevant abnormalities because the CAD mark may be on the study?

Should the ground truth include all clinically significant abnormalities seen on either the chest X-ray or the CT scan as part of the analysis?

DR. CARRINO: I think that, if the intended use of the device is to detect pulmonary nodules and it is used as a second reader paradigm, then you don't need to consider those other findings, because that would have been done as part of their routine initial assessment.

CHAIRMAN GLASSMAN: I think that is true, but if there is a significant risk that off-label use will happen, and if there is a significant risk that with concurrent reading there may be a patient risk, then do we need to consider that or does the labeling simply have to be so incredibly strong that -- and the educational process so incredibly strong that we think that the vast majority of people will use this in the intended way?

DR. STEIER: Those are two big "ifs" there, and I think what we have seen from the mammography or

colon experiences that it is somewhat likely that it will be used concurrently and for all types of things related to the lungs.

So I would be very hesitant not to consider its ability to pick up other pulmonary pathology as well.

CHAIRMAN GLASSMAN: Dr. D'Orsi?

DR. D'ORSI: I agree with that. In slide 23 that was presented to us, it says: Unlike mammography and CTC which are performed to identify a single lesion, both chest X-ray and CT are used to diagnose various chest lesions.

If that is true, then you have to put everything in there that this marks whether it is in the lung or not.

DR. CARRINO: I think we are misunderstanding, maybe confusing the part that's not -- the device is not being used to detect other things. The question is whether the observer is going to be missing those findings because of a satisfaction of search error or not doing their routine observation.

CHAIRMAN GLASSMAN: Right, my intention was in the reader study portion that we will get to in a

little while. But should we collect this as part of ground truth?

DR. SAHINER: May I make a comment?

CHAIRMAN GLASSMAN: Yes.

DR. SAHINER: In the reader study, you may not be able to get a sense of whether the satisfaction of search will be a problem or not, because I believe that in the reader's study, readers will use it as intended.

So you would be asking them to detect non-nodules before the CAD is turned on, and at that level the CAD may not have any effect on their performance because they are reading in the intended paradigm.

CHAIRMAN GLASSMAN: Dr. Steier?

DR. STEIER: Except that there is already studies showing concurrent use being validated. So it seems like concurrent use is going to be --

DR. SAHINER: Yes. If concurrent use is an intended use, then of course. Yes, if concurrent use is intended.

CHAIRMAN GLASSMAN: Dr. Garra?

DR. GARRA: Creating a test, a set, and defining whether they missed a lesion because of the CAD

being on or not, I think, would be incredibly complex. That, to me, would be excessively burdensome for a manufacturer to be able to test.

The fact of the matter is that for chest -- I would never use this concurrently because usually in my practice it is not chest nodules that I am primarily after. It is usually other pathology, and the chest nodule I pick up sort of at the end.

Now the person who is reading screening chests for nodules might end up using it concurrently, but I'll tell you that creating a reasonable set that you could even compare one manufacturer to another with a vast variety of types of chest pathology that you have would be incredibly complex, and just the analysis of that is really daunting.

I think it is very burdensome to do that. I think you need to restrict it to nodule detection and really insist that it be used as a second read.

DR. CARRINO: I would just concur with that. That is part of being a radiologist and being responsible for the entire image.

CHAIRMAN GLASSMAN: Any other comments

before I try to summarize?

DR. STEIER: Just a follow-up on that. Obviously, you two guys would only use it as a second reader. But looking through what has been presented and, agai, other experience, it seems like it is going to be used concurrently. And since we don't really do screening for nodules, that is going to be unusual.

So the reader bias and the confusion because of the markings is not really the issue for me. It is really the issue of how it is going to be used in clinical practice and what can be expected of it, and how it is marketed.

It appears as though it may be used concurrently and used for more than just screening, since screening is not what we do really. It is really for extensive nodules, morphology, the number of nodules and that type of thing.

CHAIRMAN GLASSMAN: Dr. Garra?

DR. GARRA: Is there a lot of literature showing that in the chest these things are being used extensively concurrently? I'm not aware of that.

DR. STEIER: No, just what was presented

earlier, which included studies of concurrent use.

DR. GARRA: Oh, you mean for the other types of CAD?

CHAIRMAN GLASSMAN: No, for lung.

DR. GARRA: For the chest. For instance, say if I missed a clavicle fracture, you know, on a chest -- I mean, there is such a variety of things that you could have on a chest, so much opportunity for missing things just because you miss them, that -- and it has nothing to do with whether the CAD is turned on or not. I'm just thinking that it would be really a big task to sort that out, and really hard for manufacturers to deal with that.

CHAIRMAN GLASSMAN: Let me pose Dr. Garra's point to the committee in general and see if we can get to conclusion on that.

That is, that while there is concern about satisfaction of search that from a practical standpoint in the least burdensome rule or practice, that it would be incredibly difficult and burdensome to develop a dataset to include many or all of the significant other pathologies along with breast nodules, that while the

committee is concerned about the possibility, we don't see a reasonable way to practically test for that.

Is that a reasonable agreement for us all? Okay.

Let me try then to summarize what we have said again in my pre-official -- oh, Dr. Berry?

DR. BERRY: I thought you were coming back to detection, and we have been talking about diagnosis and detection. I think they are very, very different. My comments are about detection.

It is not clear whether screening for lung cancer is beneficial. We may be doing more harm than good. I am worried that CAD would increase the harm.

Simply finding is not enough. You've got to find something that is important, and whether it is important is far from clear.

From my perspective, CAD could be approved, could be cleared for screening only if it shows clearly that it does not increase the false positive rate. So I don't know if this is ground truth or air truth, but if it doesn't show compellingly that it decreases or doesn't increase the false positive rate, I don't think any

device should be cleared for screening.

DR. GARRA: Can I just ask, a false positive rate for what? For cancer or for --

DR. BERRY: So if you are not using CAD, there is some false positive level of reading the X-rays or CT. If you add CAD, I don't want to increase that false positive level.

DR. GARRA: Just for nodules, you mean?

CHAIRMAN GLASSMAN: What if, however, you not only increase the false positive level, but the true positive level? What do you do then?

DR. BERRY: You see, that's -- I mean, the question. We see true positives now, and the question is are they clinically important? Does finding them and doing treatment -- does it increase survival? And we don't have evidence yet. We may have evidence with the PLCO and the NLST and the I-ELCAP that in fact, we do improve survival, but the wealth of studies that have been done do not show that finding cancer in the lung improves survival.

CHAIRMAN GLASSMAN: Right now, I was hoping for Ms. Brogdon to raise her hand, but we are getting no

luck there. So let's keep going. Okay, fair enough.

DR. STEIER: A comment?

CHAIRMAN GLASSMAN: Yes, please.

DR. STEIER: On the screening side, yeah, how could you really approve an advanced technology for screening when screening itself has not been accepted?

CHAIRMAN GLASSMAN: This potentially, though -- I mean, if the companies come to the FDA with simply a CAD device for nodule detection, and the intended use does not specify anything other than that, then -- I'm not trying to tell the companies what to do here, certainly. That really is not my intention. But if screening is not mentioned, then the FDA responsibilities are limited to that answer -- I mean to that use -- does that change the way we look at this?

DR. LEITCH: So theoretically, if you are doing a chest X-ray on somebody for pneumonia, that's what you are going to look for, or if you are looking for congestive heart failure.

So if you did the chest X-ray for that intent, in my mind, you shouldn't turn the CAD on because then you are screening. That is what you are doing. You

are screening for pulmonary nodules then if you turned the CAD on when your indication for the chest X-ray was for pneumonia. Okay?

Then the other issue is, okay, you are getting the chest X-ray because you are following a patient who had a previous diagnosis of cancer, and you are getting annual chest X-rays on that patient, you know, looking for pulmonary metastases, let's say, or recurrence of lung cancer.

Then you are screening. And so then you turn the CAD on, and you are trying to detect pulmonary nodules. So I think, what Dr. Berry gets into, well, does that -- you know, if you get more false positives in that circumstance and there is not a benefit to having found that but yet you have the issues of doing biopsies, all these things, what is the benefit -- cost-benefit ratio there?

CHAIRMAN GLASSMAN: These are really important issues, and I think I want to come back to them because they are a little bit peripheral to the ground truth issue. I guess the issue of whether we should be dealing with this at all is one thing, but let's see if

we can define ground truth, and then go on. Then we can deal with this maybe in the -- Now, Ms. Brogdon? Thank you.

MS. BROGDON: I'm sorry to interrupt. I wanted to make a comment to what Dr. Leitch said.

The applications that we have cleared were not cleared for screening, and the situations that Dr. Leitch just described we would not call screening.

DR. LEITCH: Right. It is not screening an asymptomatic population, but it is screening in the sense of you have -- when that person comes to the study, they are not -- you are not doing it because they have pneumonia or whatever. You are looking for pulmonary nodules in that circumstance. That is what you are looking for.

So this is the case where it is going to make a difference, is to that group of -- because you wouldn't do it -- You shouldn't do the CAD on somebody you are looking for pneumonia, if that's not what it is designed to find.

CHAIRMAN GLASSMAN: I think the distinction here is case finding versus screening, and screening has

a specific definition of asymptomatic population for prolongation of life or changing outcome.

DR. LEITCH: I understand that. I understand that.

CHAIRMAN GLASSMAN: So I think, from a definitional standpoint -- anyway, let me try to summarize ground truth, unless -- Yes?

MS. FINKEN: Just one comment from the consumer point of view. Dr. Berry, I just would like to add, survivor statistics don't always equate to the human beings. We know that we might not increase survivor over the long term, but any break we get as human beings to try to get to maybe the next stage that isn't found yet in detection techniques and in treatment is certainly valuable.

I just want to mention, as we consider the devices, these might be the key to extending that survivorship just long enough to perhaps be there when the next step in treatment comes along. So I think it is important to look at it from that standpoint also.

CHAIRMAN GLASSMAN: One quick response, Dr. Berry, and then I want to try the summary.

DR. BERRY: Yes. I mean, I certainly agree, but we would have to have evidence that, in fact, adding something does increase survivorship, and I think I mentioned that in my comments. But that is a huge task to show.

CHAIRMAN GLASSMAN: Okay, let me try to summarize ground truth.

Ground truth will be defined by an expert panel -- this is for detection and diagnosis -- expert panel, using the full knowledge that they have, including pathology, follow-up CT scan for an abnormal chest X-ray, PET scan or follow-up CT scan for an abnormal CT scan, that a full reading of the study for non-nodule ancillary findings will not be necessary. Does that accurately reflect the ground truth portion of our discussion? Dr. D'Orsi?

DR. D'ORSI: May I just make a suggestion to the FDA? Can you please let us know the criteria that you have passed prior similar devices? The diagnostic thing was like a bombshell to me anyway, and I was thinking in the detection mode all the time.

If I had known that you had cleared a prior

device for detection and not for screening, I would have had another mindset.

MS. BROGDON: Are you asking what are the approved and cleared indications for use for lung CADs and for chest CADs?

DR. D'ORSI: Right.

MS. BROGDON: Okay, we will have to compile that. I think we have that here. It will just take us a few minutes.

CHAIRMAN GLASSMAN: Okay, let's continue on then. Ms. Brogdon, that is our answer for question L1. Is that satisfactory?

MS. BROGDON: Yes. Thank you.

CHAIRMAN GLASSMAN: Okay. Question L2 will sound familiar to you. The cast has changed.

Please discuss the role of standalone performance testing in the clinical evaluation of lung CAD devices.

If you believe standalone testing should be requested in the evaluation of these devices, please provide your recommendations or comments on whether certain substrata (nodule size, shape, pathology,

location; co-morbidities; CT dose and imaging protocol; or others) should be considered in device testing and labeling.

If you believe that there are specific situations where standalone performance testing may not be important, please comment on what those might be.

Let's get started with standalone testing. First off, let's go to b. That always seems to have the least discussion for us.

Is there anybody who believes that standalone testing in any scenario is unnecessary? Something may come up as we have our discussion. We can certainly revisit this, but at first glance, do we believe that standalone testing is important?

Okay, no more comments on that except nodding of the heads as yes. Okay.

Then let's go to a. What kind of standalone testing should we do?

DR. STEIER: Can I just add a comment to that? Slide 31 says that: Unlike CAD devices for mammography and CTC which are intended to detect the only disease revealed, CAD devices for chest X-ray and chest

CT are intended to detect only one of numerous diseases and conditions that may be revealed.

Then slide 32 says: One example is solitary pulmonary nodules. But certainly not the only.

CHAIRMAN GLASSMAN: I'm assuming, despite that slide, that we are talking about CAD for the detection and diagnosis of pulmonary nodules, be they solitary or multiple. I think some of the other issues may be confounding, but unless somebody has another opinion --

DR. SAHINER: I do have a question, actually. For example, there can be devices to detect pulmonary embolism on CT scans. Are we discussing those, too, or are we only discussing nodules?

CHAIRMAN GLASSMAN: What is the intention of the question, Ms. Brogdon?

MS. BROGDON: I think we would like the Panel to discuss this as broadly as possible. If there are things like pulmonary embolism that you want to defer to the next session, which is future devices or other, you are welcome to do that. But we would like a broad discussion somewhere.

CHAIRMAN GLASSMAN: Can we hold -- oh, let me ask. Other than pulmonary embolism, are there any other pulmonary issues for CAD that we need to deal with?

DR. TOURASSI: Well, there have been other attempts for interstitial lung disease, stuff like that. So --

CHAIRMAN GLASSMAN: Can we hold this until the afternoon session on sort of future or other, because I think we are going to have enough to talk about with nodules right now.

DR. STEIER: The only thing is, though, as you consider what datasets to use and how to do your standalone testing and all that, I think it is important to know if we are just talking about solitary pulmonary nodule or if we are talking to a whole host of diseases that can be detected by CAD. And in the spirit of the broad approach, which has been mentioned, it would seem like we would have to take into account all those things.

DR. GARRA: Just looking at these descriptions -- I mean, shape, size, boundaries -- that would apply to a number of different pathologies, not just nodules. So I think we might be able to encompass

them.

CHAIRMAN GLASSMAN: Including interstitial lung disease?

DR. GARRA: Well, it does have the structure, shape, and patterns.

CHAIRMAN GLASSMAN: Okay. Then let's take them altogether, and let's see how that goes. And, if we have to break one out, we can do that.

So standalone testing: enrichment, non-enrichment, stress testing, same issues we have been over time and time again, co-morbidities -- where are people coming from? Dr. Garra?

DR. GARRA: Are we still on this question about substrata and not enrichment or did we move to enrichment?

CHAIRMAN GLASSMAN: We are talking about standalone testing in general.

DR. GARRA: Okay. Well, we already answered b. I think that everybody sort of is in agreement that we need standalone testing, and I think there does need to be substrata. But it is going to be hard with the number of pathologies to actually give a list right at

this point in time, but size, shape, density would apply to almost all pulmonary pathologies, and you might have to put in a pattern/texture measure.

I would not include CT dose, imaging protocol. I think the manufacturer should define what they consider appropriate parameters for their device and specify that these are the parameters you should be using when you perform the study that is going to be evaluated, rather than having the FDA evaluate that separately.

I think the manufacturer should make those recommendations, and that should part of the labeling that should be followed.

CHAIRMAN GLASSMAN: Other comments about -- Yes?

DR. SAHINER: I think co-morbidities is also important because the performance of CAD may be different, whether you are looking at the otherwise healthy lung or there are other diseases.

CHAIRMAN GLASSMAN: What co-morbidities should we include, without making this unduly burdensome?

DR. GARRA: That is the problem with adding co-morbidities. It could be an endless list, but the

common ones perhaps: pulmonary edema, heart failure, interstitial lung disease.

Frequently, you are asked to detect a pathology with those as underlying problems. So those are two that I can think of offhand.

DR. STEIER: Emphysema, pneumonia.

CHAIRMAN GLASSMAN: So we don't want this list to get too long, and some of these, like pneumonia and edema, will have in many instances, similar appearance. So they maybe could be put in one, which is air space consolidation as opposed to interstitial disease.

DR. GARRA: We can just recommend this as an example, a list of examples that should be included.

CHAIRMAN GLASSMAN: Right, okay.

DR. LIN: Could I ask you? Won't we end up needing thousands of patients who will be stratified by all these other things?

CHAIRMAN GLASSMAN: Well, I think the answer to that may be it depends on the intended use. If the application covers the entire spectrum of pulmonary disease, then the study to validate it would have to be

very large.

If, on the other hand, the application was for pulmonary embolus or lung nodule, the study size as determined by the relevant statistics would be much smaller. So I think it is in some way driven by the companies initially.

Dr. D'Orsi?

DR. D'ORSI: I'm getting more and more confused which is normal for an old man. But are these CADs meant for diagnosis of everything or, going back to what Dr. Naidich says, it will be restricted for its use? If it is restricted for its use, are we to take that into consideration for these standalone testing substrata?

I mean, if it's sort of nobody would put this on dealing with pneumonia, then perhaps we may not need the findings of pneumonitis. I'm just totally confused now as to where this is going. Is it diagnosis of cancer, diagnosis of pneumonia, additional nodules in pneumonia, interstitial disease, pulmonary embolism? I don't know where it is going.

CHAIRMAN GLASSMAN: I think the answer is it

depends on what the application is. What we are trying to do is broadly define the issues for pulmonary disease. If the application is just for a nodule, I think we would agree, or at least I think we would, that some element of co-morbidity analysis would be important. The same for pulmonary embolus.

On the other hand, if the application is for everything, detection, diagnosis and multiple diseases, then that would be a significant clinical first standalone and then reader study to cover statistically all of those things.

So I think we are trying to fit one size, knowing that the pie is going to be sliced up, depending on which indication or indications. Ms. Brogdon?

MS. BROGDON: I don't want to detract from what you just said, but we do have the answer to Dr. D'Orsi's earlier question about what indications for use are cleared or are approved.

CHAIRMAN GLASSMAN: Thank you.

MS. BROGDON: And Dr. Petrick will describe those.

DR. PETRICK: So these are generalizations

of the indications for use. They are in your Panel pack on pages 66 and 67, and there is a summary on page 65. But for the X-ray, the indications for the chest X-ray CAD states that the device identifies and marks regions of interest on digital frontal chest radiographs.

It identifies features associated with solitary pulmonary nodules from 9 to 30 millimeters in size, which could represent early stage lung cancer. The device is intended for use as an aid only after the physician has performed an initial interpretation of the radiograph.

So that is for the chest X-ray. For the CT lung, two CT lung CAD devices -- sorry. I'll just start here.

The indications for these devices state that these devices assist radiologists in the detection of solid pulmonary nodules during review of the multi-slice CT scans of the chest.

They are intended to be used as an adjunct to alert the radiologist to readings of interest that may have been initially overlooked. They are intended to be used as second readers after the radiologist has

completed his or her initial read.

So these are the indications for devices currently marketed.

DR. D'ORSI: So it is focused on pulmonary nodules from what I am hearing.

DR. PETRICK: The devices that are currently marketed are focused on pulmonary nodules.

DR. D'ORSI: And now there is a paradigm change to enlarge that or are we going to decide that?

CHAIRMAN GLASSMAN: That is what we are talking about.

DR. D'ORSI: Okay.

DR. BERRY: So I don't understand.

CHAIRMAN GLASSMAN: Oh, okay.

DR. BERRY: As is typical, and being gray hair as well.

CHAIRMAN GLASSMAN: I have no hair. Maybe it's easier that way.

DR. BERRY: So, Dr. Petrick, that doesn't seem to distinguish between detection and diagnosis, and is it -- does that then allow for, in a screening setting, using it to identify the nodules?

DR. PETRICK: So these devices were not cleared for a particular screening -- for a screening for lung cancer indication.

DR. BERRY: They are not cleared for screening?

DR. PETRICK: They are not cleared for a screening lung cancer indication. Right, so they are cleared for use on CTs or radiographs collected for normal clinical practice but not for a screening lung cancer--

DR. BERRY: Suppose the screening CT identifies what the doctor feels are nodules. Could then they use CAD for that?

DR. PETRICK: So they identify lung -- it is supposed to be used to identify -- help to identify lung nodules, but not -- they haven't specifically been cleared for that screening for lung cancer.

So if the study is only for screening for lung cancer, then the indications aren't included in those exams.

DR. BERRY: Okay, thank you.

DR. PETRICK: To clarify, these are

detection devices only. Although your discussion is broader, these particular indications are for detection devices. They just prompt location for nodules.

DR. STEIER: So these are for the currently approved, but there are others that will need to be looked at that will be expanding their scope?

CHAIRMAN GLASSMAN: That is why we are here.

DR. PETRICK: That's right.

CHAIRMAN GLASSMAN: That's our job, is to make recommendations for what to do in the future if there are expanded questions that are asked, expanded indications.

So standalone testing -- any other comments? Dr. Garra?

DR. GARRA: I just wanted to make one comment. Their current indications don't talk about co-morbidities and how they might affect the detection process. For instance, if a person does have bilateral mass and pleural effusions, common sense would indicate that you wouldn't try to detect nodules on a chest X-ray using that. But normally, that would be stated in the -- I think that really probably should be stated in the

labeling that it is to be used in the absence of other extensive lung disease or something if the testing was done in otherwise normal lung.

That gets around this problem of having to deal with thousands of co-morbidities and thousands of cases. If people label their device to be used not in the presence of advanced interstitial lung disease, then they don't have to test that substrata. Does that make sense? It limits what people have to test, depending on what their indications are.

CHAIRMAN GLASSMAN: Right. Dr. Dodd?

DR. DODD: This is unrelated to Dr. Garra's question. I am still struggling here with the labeling. Could somebody please help me understand? If you are not using this for screening, when are you going to use something to just detect solitary pulmonary nodules? I'm confused.

DR. STEIER: Perhaps if a patient had a mass that was already known, and you wanted to see if it was metastatic, something like that, if it had recurred, if the patient had been treated, that type of thing, if it was related to another primary carcinoma.

I mean, there may be several but certainly not to the degree that if it was approved for screening, you would use it.

CHAIRMAN GLASSMAN: On the other hand, there may be an occasional off-label use of this technology for occasional lung cancer screening, particularly CTs, but also chest X-rays that sometimes creep into clinical practice.

DR. ROSENBERG: I would assume also that normal chest X-rays, people miss pulmonary nodules. So the intent is to improve detection of nodules just in general practice. Is that not correct?

DR. GARRA: I would like to comment that that's not screening. I mean, if you get a chest X-ray for something else, standard of care is that the radiologist is responsible for picking up any pulmonary nodules there. That is standard of care and, I think, would be appropriate use of the CAD.

CHAIRMAN GLASSMAN: And for any lung/chest CT scan, you are responsible for finding nodules even if that is not the indication.

DR. STEIER: Right, the other side of the

coin is no insurance company is going to pay for everybody to have a chest X-ray and a CAD scan just in case there is a nodule.

CHAIRMAN GLASSMAN: And that is not -- We don't talk about money. So --

DR. STEIER: It would not benefit the patient -- I'll phrase it differently. It would not benefit the patient to do a chest X-ray and CAD scan on every patient looking just in case they had a nodule.

DR. GARRA: Yes, I'm talking about chest X-ray performed for other reasons, other reasons that may not be indicated either.

CHAIRMAN GLASSMAN: Okay. Let's nail down standalone testing. Any other comments before I do my wizardry here, and see if I can thread the needle for everybody? It's going to be a little harder on this one.

Okay. Part b., we all agree. Nobody has changed their mind. Standalone testing is necessary.

Our recommendations on whether certain substrata and pathology in co-morbidities: The answer is yes, and it is going to be dependent on the intended use of the device.

For example, for a pulmonary nodule, different nodule sizes should be available from, what, 5 millimeters to 30 millimeters. That's a question for the Panel. Let me finish the paragraph, and then we can come back to that number.

Different shape nodules, smooth, calcified, and irregular, different densities from ground glass through mixed to solid, all would need to be in the mix, and there should be some co-morbidity such as air space consolidation, mild to moderate interstitial lung disease, and emphysema included in the patient mix.

Let me throw that statement out to the Panel. The main question is nodule size to me. Did I pick reasonable numbers or would you rather have other numbers? I think we need to come up with something concrete here. Comments?

DR. SAHINER: I think the Fleischer Society guidelines is that if it is less than 4 millimeters, then it is follow-up CT at 12 months. I don't know what happens between 4 millimeters and 5 millimeters.

CHAIRMAN GLASSMAN: Well, 5 is what is above 4. So it would kind of fit. That's the extent of my

statistical knowledge here. You heard it.

DR. GARRA: So what does it say for less than 4?

DR. SAHINER: Less than or equal.

CHAIRMAN GLASSMAN: I think less than 4 is ignored. It's like colon polyps, unless it's high risk.

So for the performance of CAD, I think 5 to 30 is reasonable for either -- well, for CT. For chest X-ray, what do you think, 10 to 30? Ten to 30 is reasonable for chest X-ray, 5 to 30 for CT.

Now for other indications such as pulmonary embolus or interstitial lung disease, I think at that time the agency will need to develop criteria, specific criteria based on the indication, and I'm not sure in the time that we have, anything else that we will be able to come up with such specifics for those other intended uses that may come forward.

MS. BROGDON: Let me just ask the staff if they have any specific questions about these other potential indications.

No comment. Thank you.

DR. GARRA: I would like to make one comment

about maybe making them a size smaller. For a solitary nodule, perhaps 10 is okay on a chest X-ray, but if you want to detect multiple nodules such as might be a miliary TB or something, I think that there is a use for having smaller nodules being detected in that situation. They can be subtle, and you might go past a bunch of them, small ones scattered out in the lung, early fungal infection, things like that.

I think that would be helpful. I don't know that there is any downside by including smaller nodules.

CHAIRMAN GLASSMAN: I think -- I don't know. I mean, the only downside I see is the number of false positives will become astronomical, I think. What does the Panel think?

DR. CARRINO: I agree. I think that bringing it down too low would increase the false positives, and we already have a guideline from the Fleischer Society. Stick with that.

DR. GARRA: Well, then we can at least go to 4, right?

CHAIRMAN GLASSMAN: Okay, four.

DR. CARRINO: Okay. You got me down to 4.

CHAIRMAN GLASSMAN: Okay, 4. Now in terms -- there is another piece here, and that was for standalone testing, was the issue of phantom testing -- I think we all agreed that, if the CAD package includes growth and measurement, that phantom testing to validate those functions would also be part of the standalone testing. We agreed with that?

So that is question L2. Can we go on to L3, please? You've heard this one before, too, and it is really our difficult one every time.

Please discuss the role of reader performance testing in the clinical evaluation of lung CAD devices.

If you believe that performance testing should be considered, please provide your comments or recommendations on primary endpoints and corresponding clinically significant effect sizes, and again comment on ROC analysis, the merits of per lesion, per region and/or per patient endpoints in the assessment of endpoints and whether reading time should be assessed; and

b. If you believe that there are specific situations where reader performance testing may not be

necessary, please comment on what those might be.

I would like to take b. first. What do you think of not requiring reader performance testing on the measurement package of a CAD device, if one exists, or does that need to be revalidated at the reader level?

DR. GARRA: No, I agree. I don't think a reader can do that.

CHAIRMAN GLASSMAN: Is that agreeable to everybody? We can come back to it. We are not done yet. Dr. D'Orsi?

DR. D'ORSI: Can I just ask Dr. Garra a question? How valid is the measurement over time, vis à vis registration to the exact area of the mass which you would have on a phantom? Is that something to consider or not?

DR. GARRA: I think that can be an issue, but I think that the packages I have seen are not chest X-ray packages, but CT packages. They use the original data. They use the original dataset that on a modern scanner may be .4 millimeter sizes, and they find through the slices the same slice and everything, and there is no way a human observer is going to be able to even approach

that kind of performance.

So I don't see how a human observer study is going to help there.

CHAIRMAN GLASSMAN: Any other comments about this, having to do with measurement or anything else?

DR. SPINDELL: Just to go back to the same question we had on the other ones, for simple, minor algorithm changes would we require reader performance testing again?

CHAIRMAN GLASSMAN: Dr. Berry?

DR. BERRY: Yes, the way you asked the question had so many negatives, I'm not sure whether to answer yes or no, but I don't understand how you can consider a submission that wouldn't involve the role of the reader, and what does CAD add to the normal circumstance.

So I'm not sure what we are talking about here, but the reader is essential in every submission.

CHAIRMAN GLASSMAN: We are talking specifically about the measurement piece of a CAD detection or diagnosis. It would be at that point probably more diagnostic system. Just the part that says

how big is it? Has it grown? Is it a different shape than it was the last time?

DR. BERRY: So that could be standalone.

CHAIRMAN GLASSMAN: No, it would be standalone.

DR. BERRY: Oh okay, all right.

CHAIRMAN GLASSMAN: But the question of whether a human reader needs to be involved in that small part of the equation was the question that I posed.

DR. BERRY: I see. I see. Okay, I understand.

CHAIRMAN GLASSMAN: Dr. Bourland, or Dr. Dodd first, then Dr. Bourland.

DR. DODD: Right, so I don't want to address the issue of a reader study in that setting, but I wonder, since there are limitations to the phantom studies, is there someplace in between where you could take a set of data for lesions that have been monitored over time for which you know there is no significant change and another subset of lesions -- nodules, rather, that you know have grown, to establish whether that accurately characterizes some sense of growth of those

because this registration is an issue, and phantoms may give you an overly optimistic estimate of the performance in practice?

CHAIRMAN GLASSMAN: Dr. Bourland?

DR. BOURLAND: So I think it is hard to predict for this quantitative side whether the role of the user, the role of the reader. Is that a training role? Is it an implementation role? I don't know, but I think there is value -- there could be great value in terms of a reader impact on some of these results.

I understand the idea of, if I've got an area defined on five different slices of thickness 5 millimeters each, I can multiply them by five and -- you know, et cetera. Perhaps that part of the computation does not need to be tested again, but I think the user connection with how that all works is actually a very critical piece and would thus deserve some amount of testing.

CHAIRMAN GLASSMAN: Is that the general feeling of the group? So for b. then, there would be some reader testing of the measurement package, but for minor algorithm changes, whatever the definition of minor

turns out to be, re-reader testing would not be necessary. Can we agree to that? I guess so.

Okay, let's go back to the hard part then, to a. -- appropriate primary endpoints and corresponding clinically significant effect sizes and ROC curves. Dr. Berry, do you want to take a shot at that first?

DR. BERRY: No.

CHAIRMAN GLASSMAN: Dr. Abbey or Dr. Dodd? Somebody? You've just pointed to him, and he nodded at you. Okay, Dr. Dodd's light is on.

DR. DODD: I think many of the issues are similar to what we have already discussed. So I don't have much to add except I want to emphasize the importance if we are talking about detection or screening for lung nodules, there should be a much greater emphasis on specificity.

CHAIRMAN GLASSMAN: Do you think that -- oh, Dr. Abbey.

DR. ABBEY: I was just going to echo that, but I don't have anything to add.

CHAIRMAN GLASSMAN: What about -- do you think that the ROC analysis or FROC or just sensitivity

and specificity -- I know we have touched this before, but specifically for lung CAD nodules and other things, what is the most appropriate statistical way to look at that?

DR. TOURASSI: Again, it depends if it is diagnosis or detection. For diagnosis, an ROC study is certainly appropriate if we are looking at the per lesion analysis; and if it detection, FROC or JROC, of course. But this is an open-ended discussion. We don't know what this is we are talking about, and if it is detection or diagnosis.

CHAIRMAN GLASSMAN: That's the hard part. So in general, are we agreed for ROC analysis for diagnosis and for detection FROC or JROC would be the preferable statistical tools?

What about endpoints and clinically significant effect sizes? I know it is hard when we don't know which disease, but how about for nodules, which is probably the most likely submission? Is there any statistical sense by anybody that there are any numbers we can throw out?

Let's move on. Maybe we will come back to

this one.

The merits of per lesion, per region and/or per patient endpoints in the assessment of endpoints? Nodules and then others? Why don't we separate it out that way, because the answer may be different? What about for nodules? Per lesion, again for detection. We are looking for lesions.

Which of these endpoints is most useful? Dr. Dodd?

DR. DODD: Well, I have certainly seen per lesion and per region both used. I think either one is acceptable.

CHAIRMAN GLASSMAN: If we were to use -- Well, I guess for the FROC, we could use either. Is that right? We could use per lesion for FROC, but not ROC? Do I have that right from yesterday?

DR. GARRA: Well, FROC typically gives you performances, false positives per image. So it is really kind of a per image metric.

DR. DODD: If you are doing per region, you could do an ROC analysis, accounting for the correlation.

CHAIRMAN GLASSMAN: Okay. Now what about

some of these other things pulmonary embolus, interstitial lung disease, where we don't -- well, for interstitial lung disease, we clearly don't have a single lesion. It is a very diffuse process, and it is sort of present/absent and level of confidence, but it is -- would we use region where region is the whole lung? Would that be the appropriate way to approach that specific issue or are we getting ahead of ourselves here?

DR. STEIER: You might be able to get by with a generic statement, "appropriate measures as new indications are proposed," instead of trying to anticipate a whole host.

CHAIRMAN GLASSMAN: Okay. Dr. Garra?

DR. GARRA: I agree with that comment, except you can give a couple of examples. For instance, per patient or per lung could be used for diffuse lung disease, and for nodules you could use per region or per lesion as examples.

CHAIRMAN GLASSMAN: Is that -- any other comments about -- Dr. Berry?

DR. BERRY: We don't want to lose sight of the forest for the trees. The focus has to be in doing

everything we do on the patient. It is true that we may want to analyze the lesions or the regions, and we may want to do a hierarchical analysis, but the focus has to be on what impact we are having on the patient.

DR. SAHINER: May I add something? I agree with that, and as was previously said, if you detect the 16th nodule on a patient with 15 nodules, if you are doing per lesion analysis, it would count as an additional detection, but for patient management it may not have any effect.

CHAIRMAN GLASSMAN: However, I'll say what I said yesterday, and then see how everybody feels. If you are talking about lung CT as a clinical device, per patient makes a lot of sense. To me, not being a statistician, we are talking about a CAD device who is at least in the first cut before diagnosis, is simply a lesion finder, and it may make a difference whether you find one lesion, two lesions, three lesions or the 16th lesion if you are talking about detection efficiency and accuracy.

DR. SAHINER: May I?

CHAIRMAN GLASSMAN: Yes.

DR. SAHINER: That's correct, but in a reader study if a reader finds 15 nodules, he or she may be much less vigilant to find the 16th nodule because they already have some kind of analysis for the patient. So in a reader study, it may skew things if the patient has too many nodules.

CHAIRMAN GLASSMAN: On the other hand, if the patient has two nodules and the reader or CAD missed one, then that becomes much more significant.

DR. LEITCH: Again, I think for the application of the technique in patient populations, for what percentage of patients does it make a difference that they had that test done, and is that difference a good difference or a bad difference? I mean, that's why you have to do the patient part of it.

Your example is an example where it would make a difference in the patient, you know, if they have a solitary versus multiple nodules. That would make a difference for the patient, but the other issue of 15 versus 16 doesn't make a difference to the patient.

So any technique that is applied, the bottom line is does it help the individual patients in the whole

population that you are screening or applying the test -- not screening but applying the test -- or does it cost harm to what number of patients?

CHAIRMAN GLASSMAN: Dr. Garra?

DR. GARRA: We can pick it apart endlessly, because it depends on the population. For instance, if there is a patient with 15 nodules and they want to know has his metastatic disease progressed or not, and I see 17, then I call it progression. That has a huge impact on that patient. If it is granulomatous disease, it doesn't have an impact. It depends entirely on the patient, but what I care about when I am using a tool is raw performance. Is it going to help me find those additional nodules that are going to make all the difference in the world to some patients and may have no difference on another? That is my responsibility as a radiologist to make that call.

DR. LEITCH: But is this test -- are they saying because this is another issue of using these techniques or if they are comparing to the last study? So how good is -- you know, if you are asking the sequential question, if you applied CAD, would you be

more likely to accurately assess --- you know, the first time you did it there were five nodules, and then the second time you did it there were six, and one of them increased 2 millimeters? Would CAD pick that up better than the doctor would pick that up?

So is that what the application is, or not?

CHAIRMAN GLASSMAN: Oh, there are multiple applications.

DR. LEITCH: You know that there is a sequential look. Can the CAD do that?

DR. CARRINO: I would just emphasize, I think Dr. Garra has put a balanced perspective and stated it very precisely and concisely, and that is the paradigm that, I think most radiologists would use and that is I think the intended use for these devices, and that is how they would probably be labeled.

CHAIRMAN GLASSMAN: Let me remind everybody, if my memory is still with me, that yesterday we picked -- and this is another issue, but we picked per lesion or per region when we entered this question about other CAD devices, and I'm a little uncomfortable --

DR. BERRY: That was standalone.

CHAIRMAN GLASSMAN: Oh, this is reader, right. Thank you.

DR. TOURASSI: Actually, you are right. For the reader performance, we also agreed because we had that argument, but yesterday the colon cancer application is very different because there appears to be viability among clinicians as to what is a proper patient management depending on the size, 5, 6, 8, 10, whatever. But here for the lung cancer, the per patient analysis for the reader observer study makes more sense.

I agree with Dr. Sahiner, with Dr. Berry, because in the end it will help you find that 17th nodule which -- as you say, it is important for you. Then the CAD system will help you change your decision regarding patient management.

So that will be reflected in the sensitivity and specificity as measured on a per patient basis. So I don't see why the extra burden of the per lesion analysis for the reader observer study. In the end, it will impact your patient management decision for the patient, even if it's the second nodule or the 17th nodule.

CHAIRMAN GLASSMAN: Dr. Garra?

DR. GARRA: Well, in some patients, it would, not in others. So --

DR. TOURASSI: But that is the point. In the end, if CAD makes a difference, it will be reflected there. It doesn't matter which lesion number it is.

CHAIRMAN GLASSMAN: Go ahead, and then I will have a comment.

DR. SAHINER: For nodule detection, can we agree on a, for example, maximum number of nodules that would make sense? I know that, for example, for the LIDC study or some other studies there was mention of six nodules or more. So if you limit for the -- I'm not sure, but if we can come up with a number for the number of nodules by which they wouldn't be counted as additional picks by the computer.

DR. TOURASSI: Well rationally, it makes sense. But there is, of course, this debate among the clinicians. Can we fix a number? How? There is no proof or whatever.

CHAIRMAN GLASSMAN: Let me ask the Panel. The added burden of collecting data on a per lesion and reporting data on a per lesion and a per patient rather

than one or the other -- since we are somewhat split here, if the burden is not significant -- that is, to do per patient, you have to collect per lesion anyway, then can we agree that both endpoints are desirable, rather than come up with just one?

First off, is there a burden difference that is significant?

DR. ABBEY: That was going to be, I guess -- I was going to ask the same question, but I think it does become a burden if the statistical power of the two studies is very different. Then you have to power to the weakest of the two if you require both endpoints.

So I would think that a per patient study is actually more burdensome in the sense that you are going to require more cases and more reads; whereas a per lesion, if you have 17 lesions in an image, you've got 17 responses that you are going to compress into one per patient.

CHAIRMAN GLASSMAN: Dr. Berry?

DR. BERRY: I think it is more burdensome. I agree, but it is an essential aspect. You've just got to look at the per patient.

To the point of Dr. Sahiner and to Dr. Garra's example of the 17th nodule in a diagnostic setting where you have a cancer patient and the RECIST criteria indicate that a new nodule, a new lesion is a progression, that is critical. So you can't put a number on them, but I think that the focus is on the patient. The analysis may be at the lesion level, and it is the more burdensome part, but it is appropriate.

The least burdensome doesn't mean that you are going to throw out the baby with the bath water.

CHAIRMAN GLASSMAN: Comment? Dr. Bourland?

DR. BOURLAND: I just have a comment on left and right, and that I think that is important. If we have per lesion that also tells location, but in some cases disease might be confined to one lung and knowing the contralateral side was somehow free or freer of disease might actually be very helpful.

For instance, radiation beams -- you know, in general we would say we want to stay out of the contralateral lung anyway but might actually have some impact on care.

CHAIRMAN GLASSMAN: Dr. Dodd?

DR. DODD: So I just want to emphasize that I, too, think per patient analysis is very relevant and critical for screening because what follows that is probably a diagnostic CT scan. But I have also assumed in any of these per region analyses, per lesion analyses, that you are going to do the per patient analysis as well, because those things begin to put together a picture of what is actually going on.

When you get to the diagnostic CTs, I'm a little more uncertain because those are actually -- the location is important as far as I understand for the diagnostic workup, and the per patient thing is still important, but you may want to at that point look at the per lesion and the location information.

CHAIRMAN GLASSMAN: Do we feel that location needs to be varied from hilar to peripheral and lower lobe and upper lobe, as well as size for the reader study portion of this?

DR. STEIER: Yes, it has implications for staging, treatment, et cetera. Of course, you need that.

CHAIRMAN GLASSMAN: For the ease of detection. Dr. D'Orsi?

DR. D'ORSI: This is a new paradigm that we are looking at from everything else we looked at. We are mixing detection and diagnosis here. So I think what you said is very clear.

For the detection phase per lesion is fine. I think on the diagnostic phase, you will focus on the patient to see exactly what is going on as far as affecting the patient. Diagnosis is going to affect the patient.

So as you said, if you are going to collect per lesion, you automatically are going to deal with per patient effects, if this is a dual detection and diagnosis. I still have problems with that.

CHAIRMAN GLASSMAN: For diagnosis, I presume that the ROC type analysis will be critical, or some variant of it, and there certainly, I think, per patient has a lot -- if we are dealing with nodules and lung cancer, certainly has a lot.

I'm still not sure that we have come to a consensus. What about per patient versus per lesion or region?

DR. BERRY: I just wanted to get a -- when

you do per patient, you are throwing out all location -- You don't have to. To do per patient, you can specify for a patient criterion that an increase in number or size nodules or change in position or whatever are significant.

If you are capturing that anyway, why not do per lesion because you do get the per patient information anyway?

DR. TOURASSI: I guess my question -- I was trying to figure out what is the most burdensome for the detection task, in particular, to ask the reader to mark, to give a rating pretty much for every nodule they have identified or to ask them to read every case and tell you, yes, I will send them for a follow-up CT; no, fine? But for that paradigm, as Dr. Abbey said, we will need a lot more cases.

So what is the least burdensome of the two, and what makes, of course, more sense in the clinical significance?

DR. GARRA: Well, I'm thinking if we don't have lesion location, we are not going to know anything about location in the lung. We are going to be missing a

lot of information that is important from a treatment standpoint, you know, in addition to diagnosis.

So you are going to have to capture that anyway somehow. So if you are going to do it anyway, then doing -- I definitely agree that per patient is very important, but when you do per lesion, you will have the per patient information, too.

CHAIRMAN GLASSMAN: And so we shouldn't throw that out. Dr. Sahiner?

DR. SAHINER: I just agree with Dr. Garra that per lesion is actually less burdensome. And although per patient is important, I think per lesion is also important.

CHAIRMAN GLASSMAN: Certainly, per lesion would be important if we are trying to figure out false negatives and geographic areas near the hilum and things like that, and also false positives. If we just deal with the patient, we lose that ability.

So if the per patient is more burdensome but it is important, do we think that -- I think we need per lesion in my mind, and I will throw this out, and everybody can -- you know, when I do my pre-speech.

Given the burden of per patient and given the importance of per patient, does one outweigh the other, and should we come down on the side of both per lesion and per patient or just per lesion for reader studies? Dr. Berry?

DR. BERRY: Dr. Glassman, I don't think there is much disagreement, actually. I mean, when I stated my position, I said it will be essential to evaluate the lesions within the patients.

So I am looking sort of top down, and Dr. Garra is looking bottom up, but we come to the same conclusion. Both are important, and there is no way -- and I think even with Dr. Garra's position, I don't think there is any way that you get rid of the burden of looking at the patient.

So exactly how you state it probably doesn't matter much. Both are important. My own perspective is that in the statistical section you would write things in terms of the patient, but a fundamental part of that will be evaluating the lesions within the patient.

CHAIRMAN GLASSMAN: Okay, any other comments? Dr. Dodd?

DR. DODD: I just want to agree with that. I think the study should be powered for the per patient analysis.

CHAIRMAN GLASSMAN: Okay, I think that is very important. Let me give this a try.

Performance testing is definitely critical for reader performance testing. Primary endpoints and clinically significant effect sizes: the committee really has no hard recommendations -- it will depend in part on the intended use whether it is diagnosis or detection, and what we are diagnosing or detecting.

ROC analysis is felt to be very important, particularly on the diagnostic end with JROC or FROC on the detection end. We think that per patient endpoint is very important, and that the study should be powered to give that information.

Along with that, per lesion information to evaluate for false positives and false negatives will be important.

What about reading time? Again, previously, we have said that that should be measured. Do we agree for these techniques also? Does anybody think it

shouldn't be measured? Dr. Abbey?

DR. ABBEY: I'm not sure what it tells you in a case where you've got a lot of potential other things you are looking for. How much would that -- I just have no sense of whether that measurement conveys much information.

CHAIRMAN GLASSMAN: Is that because there are so many co-morbidities that you will be evaluating that the increased incidence -- the increased time of CAD will be really not very helpful in any analysis?

DR. ABBEY: Yes, that's my concern, is that we will measure something that is so corrupted with other factors in it that it won't have much information. But it's easy to measure. It's just should it be included in the analysis or required to be included? If it's there, what do we do with it?

CHAIRMAN GLASSMAN: Yes?

DR. LEITCH: But you will be measuring the time without the CAD, too. So you could do it -- you know, the difference, not so much how long an individual exam took but just the difference of adding the CAD to it.

DR. ABBEY: I suppose, if the cases are matched, then that's fine. If the cases have different co-morbidities associated with them, the variance --

DR. LEITCH: So you are going to stratify for that anyway. Right?

DR. ABBEY: Okay.

CHAIRMAN GLASSMAN: So again, time or no time? We have one optional. Dr. Spindell?

DR. SPINDELL: Yes, and I know we had this discussion on all the previous ones as well. I understand we may want to measure it, and I understand all the socioeconomic, medical economic things, but as far as safety and efficacy, how does the reading time play into the safety and efficacy decision that the FDA will ultimately have to make?

CHAIRMAN GLASSMAN: Very good question. It doesn't unless the time is so burdensome that it causes a distraction from the co-morbidity issue.

DR. SPINDELL: And wouldn't that be measured in the reader time, the sensitivity-specificity ROC curve analysis, et cetera?

CHAIRMAN GLASSMAN: Dr. Garra?

DR. GARRA: I tend to agree with this comment, and if the time is really excessive, the market forces will eliminate that product as well.

CHAIRMAN GLASSMAN: Okay, so no time for lung CAD. Ms. Brogdon?

MS. BROGDON: Our question is whether the Panel believes this information is critical to the prospective user of the device.

CHAIRMAN GLASSMAN: Dr. Garra?

DR. GARRA: No.

CHAIRMAN GLASSMAN: Any other comments? I would concur with that. I don't think that time -- the difference here is the other two modalities -- we were talking about breast and colon CAD -- are really in screening situations where the time is much more relevant than in a diagnostic situation which is what we are talking about here. I think that is probably the reason why it is not important.

DR. STEIER: A comment? I agree. I think the issue here is not time but accuracy. So I would not -- time would not be an important factor, speaking as a non-radiologist.

CHAIRMAN GLASSMAN: But you do take those automatic EKGs, don't you?

DR. STEIER: Yes, but I read them first, always a second reader.

CHAIRMAN GLASSMAN: Okay, does that answer for L3 satisfactory for the needs of the agency?

MS. BROGDON: Yes, thank you.

CHAIRMAN GLASSMAN: Thank you. It is now 12:25. In the hopes that L4 will be brief, I'd like to do that, and then break for lunch if we can. Is that reasonable? Okay. I've gotten permission to do that from my boss here.

L4 -- let's see if we can get that one done. Please discuss whether there are other types of performance testing you believe should be considered in the evaluation of lung CAD devices.

Any other? No lights are going. Dr. Abbey?

DR. ABBEY: I will just make a quick statement that there are emerging methodologies in multi-class ROC kind of analysis that do make sense sometimes when you have co-morbidities along with it. I don't think they are ready yet to require companies to do that.

CHAIRMAN GLASSMAN: So the answer is not at this time. But if statistics change, it may change. Is that an acceptable answer to L4?

MS. BROGDON: Yes, thank you.

CHAIRMAN GLASSMAN: Thank you. Let us break for lunch. It is now 12:26. Why don't we come back at 1:26, please? See you then.

(Whereupon, the foregoing matter went off the record at 12:26 p.m.)

- - -

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

Time: 1:29 p.m.

CHAIRMAN GLASSMAN: Okay, I have called us to order. We will go back to lung question L5. We had just finished L4, and we will continue in our discussion of lung CAD and then move on to general and future issues.

The prevalence of lung cancer cases in the population having chest X-rays and chest CT is relatively low. Please provide comments on the practice of using an enriched dataset for the clinical evaluation testing discussed in 2, 3 and 4, which is the standalone and reader testing.

If you believe that an enriched dataset may be used for these evaluations, discuss what you believe to be the appropriate clinical, imaging and pathological characteristics for that database.

Please consider items such as number of patients with no nodules, single nodules, multiple nodules, range of nodule sizes and (b) if you believe that enrichment is inappropriate, please provide your reasons and whether there would be an alternative method

of assessing these devices in light of the low prevalence.

As we did before, could we discuss part b. first? Does anyone think that there is a way to evaluate lung cancer screening -- or lung cancer CAD -- I apologize -- lung cancer CAD without an enriched dataset that is anything that would be least burdensome? Any comments?

Are we in agreement that -- yes, Ms. Brogdon?

MS. BROGDON: You mentioned lung cancer.

CHAIRMAN GLASSMAN: It was mentioned in the question. Would you rather this be a broader discussion?

MS. BROGDON: Well, I'm not sure that it helped our previous discussion.

CHAIRMAN GLASSMAN: I was very happy to see it in this question, I have to admit.

MS. BROGDON: Okay, proceed.

CHAIRMAN GLASSMAN: Okay, let me broaden it just a little bit then. For other lung conditions such as pulmonary embolus or interstitial lung disease which have a relatively low prevalence, are there anything

other than enrichment that would be helpful for looking at this that would fall into the least burdensome category? Does anyone have any comments?

If not, I will take that as enrichment is appropriate.

Now let's go back to (a), if we can. Appropriate clinical imaging and pathological characteristics for -- let's first deal with lung nodule, since -- lung cancer, since that is specifically asked in the question, and depending on time, which we are not likely to have, we may spend some time on the others. But lung cancer.

Different types of pathology, different clinical settings for the database that is used to test these devices -- comments? Yes?

DR. SAHINER: For clarification, may I ask are we talking only about lung cancer or lung nodules?

CHAIRMAN GLASSMAN: The question is lung cancer, but I think -- Ms. Brogdon, can we broaden this to lung nodules or would you rather keep it as cancer?

MS. BROGDON: I think nodules would be fine.

CHAIRMAN GLASSMAN: Okay, lung nodules. So

pathological characteristics -- why don't we deal with that first. What spectrum of diseases in the lung that form nodules should be included based on their imaging appearance or their pathological basis? Adenocarcinomas?

DR. STEIER: I would guess a representative set might include a couple of lung cancer cases, a couple of cases of sarcoid TB, other types of diagnoses that can cause nodules.

CHAIRMAN GLASSMAN: And among the cancers, any subtypes that you would like to specify be included, small cell, large cell?

DR. STEIER: A representative sample, any of the above, bronchoalveolar perhaps.

CHAIRMAN GLASSMAN: Any other comments about pathologic types of things that form nodules? Yes?

DR. LEITCH: Just be sure you have some metastatic nodules as well as primary lung cancer nodules.

DR. STEIER: Sure. Even septic nodules.

CHAIRMAN GLASSMAN: Okay, so we have got a broad spectrum of nodules. What about characteristics? What about no nodules? How enriched should the sample be

for first standalone testing and then reader testing?

Let's deal with standalone first. In the past, we have said that the richer the better because it is standalone and it really -- the computer doesn't really care. Would that be consistent again here? Yes?

DR. KIM: I think you would want a proportion of completely normal cases so you know what the false positive is on the normal exam.

CHAIRMAN GLASSMAN: So -- Yes?

DR. CARRINO: For computer standalone purposes, you can have lots of cases because it's just a computer. So the number -- the total n, I think, is less relevant. You can use the big N.

CHAIRMAN GLASSMAN: So a large dataset, lots of normals, and lots of abnormals but not at the low prevalence.

What about stress -- Dr. Garra?

DR. GARRA: They can be at a high prevalence because, I mean, the system is not going to learn what the composition is.

CHAIRMAN GLASSMAN: Exactly, now what about stress -- oh, I'm sorry. Dr. Dodd?

DR. DODD: I was going to say though with regard to normals, are there things that might be confused with the lung nodules of interest that you would want to -- that are considered normal, that you would want to include in the standalone testing?

DR. STEIER: Sure, scarring and other anatomic things like that. So yes, you could have a selection of things that are commonly confused for nodules as well.

CHAIRMAN GLASSMAN: Okay.

DR. BERRY: Can I just add?

CHAIRMAN GLASSMAN: Yes, Dr. Berry.

DR. BERRY: This is -- somebody made a point earlier yesterday about the importance of doing enrichment, and if I can have a wish list, I would love to see an enrichment based on eventual outcome, you know, those lung cancers that were fatal and quickly, those lung cancers that weren't so much.

Here is an opportunity where you can take historical settings and put them in front of the CAD and see how it does. I think that this would be an exploratory setup, but the potential for identifying --

you know, what are we finding? Are we finding the really bad things or are we finding the things that we really shouldn't have found?

CHAIRMAN GLASSMAN: Other comments about that kind of stratification? Yes?

DR. SPINDELL: I would say that should be based on the indications for use and what the device is actually intended to detect. If it is just to detect nodules, I grant that that is great information, but I think that might be not the least burdensome approach on the manufacturer.

DR. BERRY: No, I meant that this would not be the registration in the indication, but it is an opportunity for the company to see what they've got.

So in the early phases of the development, these are the kinds of things that you really want to do.

CHAIRMAN GLASSMAN: Dr. Ziskin?

DR. ZISKIN: Because of the importance of false positives, I think there should be an adequate number of normals, so you can get some assessment of this.

CHAIRMAN GLASSMAN: What about if I throw

out a specific number? I hate to do this, but I am going to do it anyway. Fifty percent normals, 50 percent nodules, a mix of singles and multiples -- is that a reasonable stress that will give a good read, and some of the normals will have some variance in them that are confused with nodules. Would that give a good reading on the false negatives and the false positives?

DR. BERRY: As long as it is standalone, it doesn't matter. I mean, yes, total sample size. If it is in a reader setting, then I think it does matter, and I think you would want to use the strategy that I suggested earlier of varying the rate over the course of time.

CHAIRMAN GLASSMAN: Yes, we are just talking about standalone right now though.

DR. BERRY: So I'm not sure you should say 50 percent. I mean, just say enough of both, or something like that.

CHAIRMAN GLASSMAN: Okay, we'll get to the summary in a minute, Dr. Garra.

DR. GARRA: Just another specification for types of nodules, specifically nodules that have been

missed. You can get that from RAD peer data. You can get that from published studies that show certain areas of the lungs where nodules are missed, and make sure there is plenty of those.

CHAIRMAN GLASSMAN: Okay, so range of nodule sizes -- we kind of covered that before. Are we still 4 to 30?

DR. GARRA: It was 2 to 30, wasn't it?

CHAIRMAN GLASSMAN: No, it wasn't. Good try from Vermont. We used 4 to 30 before. Is there any sentiment to leave that the same, or to move to a different set of numbers? The same? Okay.

Now what about the reader testing, which is, obviously, a much more complicated issue? The level of enrichment will be different.

What about, though, the mix of cases? Forgetting numbers, what about types of cases? Would you change the type, or would you leave the types of cases the same and just change the prevalence because we have defined a group of cases for the standalone testing, different kinds of nodules, different kinds of non-nodules in the normal group. Would you leave that mix

the same? Dr. Dodd?

DR. DODD: I thought yesterday, with regard to reader studies, we discussed having a representative sample when it comes to a reader study and then ensuring that you have enough numbers in that representative sample for some of these.

CHAIRMAN GLASSMAN: I think all of these are representative. The mix -- I mean, now --

DR. DODD: I'm talking about the mix.

CHAIRMAN GLASSMAN: Yes.

DR. DODD: A representative mix.

CHAIRMAN GLASSMAN: Yes, the mix will be -- Okay, let's talk about mix.

DR. TOURASSI: But this is going to be more challenging now without a clearly defined population. For breast CAD and colon CAD, we were talking about screening populations. Here, what is the population?

CHAIRMAN GLASSMAN: It will be non-screening, presumably.

DR. TOURASSI: So the study characteristics, the prevalence is well defined for the non-screening population when it comes to chest X-rays and lung CTs?

DR. GARRA: Yes. For Midwest, it will be all histoplasmosis. Right?

CHAIRMAN GLASSMAN: Right, so there are some geographic differences, but I think there is data on the prevalence of different lesions in what are considered general populations in terms of different kinds of cancers.

Prevalence of TB and histo and things like that are often geographically or socioeconomically skewed in one way or another, but these things could be included in some rough reasonable mix without it being too skewed, I think, although I don't have hard numbers. Is that a reasonable thing?

DR. GARRA: I agree that you could do that. You could take, for instance, the population in the United States and take -- you could weight them geographically by the relative population of various parts of the country where we know they have higher incidence of certain things, like coccidioides or something in various parts of the country.

CHAIRMAN GLASSMAN: So again, patients with no nodules would be the majority in the reader study

because that is the majority in the population, but that there would have to be enough of these to see whether the readers perform well. Dr. Ziskin?

DR. ZISKIN: I would like to talk about the size issue of nodules be smaller than 4 millimeters. If it turned out that the algorithm used by the CAD actually would get tripped up and would actually call these small nodules as such, these would be a false positive.

How would we know that if you didn't have any very small nodules in the test set?

CHAIRMAN GLASSMAN: Anyone want to agree or disagree? Yes?

DR. LEITCH: I think that is an issue, because on chest CTs those smaller nodules are seen, and you know, when they are seen in the context of somebody who has had cancer in the past, you are never allowed to just kind of walk away. It is something you do have to pay attention to.

So because that is a sensitivity of the CT, I think you -- you know, with the CAD you've got to address that issue.

CHAIRMAN GLASSMAN: Anyone else other than

Dr. Garra, and we know what he thinks? No Brian, if you want to say something, go ahead.

DR. GARRA: Didn't we just specify that they were going to be 4 millimeter nodules in the dataset?

CHAIRMAN GLASSMAN: We just discussed it. That doesn't mean that we can't change it. Four to 30 is what we said. Yes, Dr. Dodd?

DR. DODD: I thought Dr. Ziskin was suggesting that we include some below 4 millimeters.

CHAIRMAN GLASSMAN: He was, and that's what we are discussing.

DR. DODD: That falls under my category of the nodules that are confused -- no-nodule situation that are confused as nodules.

DR. STEIER: Right, that goes with the false positives where you are going to get, you know, hundreds or thousands of small nodules less than 4 millimeters that you are going to be stuck with, try to track down, and figure out what to do with. So you would not want those. So I think you would test to try to avoid those.

CHAIRMAN GLASSMAN: Those then, though, would be included in the "normal" portion rather than the

enriched portion, but they should be included is what I am hearing. Okay?

DR. ROSENBERG: So 'no nodules' would include small nodules?

CHAIRMAN GLASSMAN: Yes, no nodules less than 4 millimeters. Normal would be no nodules 4 millimeters or greater. Dr. Abbey?

DR. ABBEY: A question of clarification in designing these studies. Should the study be powered for the individual lesion type or just for the entire -- powering for one single study, as opposed to the individual kind of lesion? So do you want to be able to make a significant claim about this size lesion, that size lesion, or do you just want to be able to say we are substantially equivalent across the entire study?

CHAIRMAN GLASSMAN: Dr. Berry?

DR. BERRY: The entire study. You can't -- this would be really burdensome to try to address individual lesions.

CHAIRMAN GLASSMAN: Very good point. Thank you. Any other comments about L5? Okay, let me try to summarize. See if you like it. If you like it, we will

go forward.

Part (b) We believe that enrichment is appropriate.

Part (a) For the standalone testing, we believe that a high prevalence of abnormals with nodules from 4 to 30 millimeters be included, and that a sufficient number of normals be included, and normal or benign encompassing nodules smaller than 4 millimeters, scarring, sequestrations, and other things -- azygos lobes we didn't mention but that would be one that could easily be confused with a nodule -- and that the nodule types come from multiple pathologies, including carcinoma and infection, sarcoid, septic potentially, and the inclusion of metastatic multiple nodules.

Does that reflect our discussion? Is that sufficient, Ms. Brogdon?

MS. BROGDON: Yes, thank you.

CHAIRMAN GLASSMAN: Thank you.

DR. GARRA: Dr. Glassman, I have a question. So how would a 2 millimeter nodule that was detected by CAD be scored then because there is really a nodule there in a case where there is really a 2 millimeter nodule

there?

CHAIRMAN GLASSMAN: Over here I am hearing false positives.

DR. GARRA: I know.

DR. D'ORSI: If you define positive one way, you have to stick to it.

DR. GARRA: I just have a problem with using non-English definitions.

DR. D'ORSI: Well, I mean, if your definition is negative is nothing or polyps less than 4, that's negative. So if you find it, that's a false positive.

CHAIRMAN GLASSMAN: I am going to leave that to the FDA.

DR. GARRA: Yes, we can leave it to the FDA. I personally disagree with that. I think that frequently a nodule when it appears at 3 millimeters on a CT in a person with metastatic -- potential metastatic disease is significant.

CHAIRMAN GLASSMAN: It may be, but we are evaluating the CAD system. So let's leave it up to FDA to make that decision when the time comes.

DR. SAHINER: If I may, if you are talking about the reader study here -- you know, the observer reading with CAD -- hopefully, they won't mistake a 2 millimeter nodule for a 4 millimeter nodule especially if they have some tools for measurement.

DR. GARRA: That's a difficult measure.

CHAIRMAN GLASSMAN: Sure, Dr. Berry.

DR. BERRY: Can I add that the -- my point yesterday was exactly on target with Dr. Garra's point today that the ROC and the sensitivity specificity are inherently binary, and this is a situation which is not inherently binary and some improvement on the ROC analysis to take into account this sort of thing, I think, would be helpful.

MS. BROGDON: They are nodding.

CHAIRMAN GLASSMAN: That will be noted in the record. I am hoping that the statisticians at the FDA know exactly what that meant. That is not your ability to state it, Dr. Berry. It is my lack of statistical knowledge that I am commenting on.

MS. BROGDON: They are nodding, yes.

DR. BERRY: Well, it's very simple. It is

positive or negative, and this is a case that is not positive or negative.

CHAIRMAN GLASSMAN: Okay, somebody will teach me more about ROCs than I know at another time.

Let's move on to L6. FDA does not specify indications for use but reviews indications for use that are requested by companies. What are the Panel's views regarding second reader versus concurrent reading of a CAD device?

How are lung CADs used clinically?

Are second reader and concurrent reading modes both relevant options and, if not, which is appropriate; and do we believe that users understand that if something is labeled for second read, that that is the way they should use it? A similar question, but a different issue.

Lung CADs used clinically -- I think we said earlier both ways. Does that reflect even in the literature? Does anybody take issue with that statement? Okay. Let's move on to the next one.

Are they both clinically relevant? There was some data this morning that showed -- that we talked

about that they were. Is that -- does everybody go along with that? It's a little different than what we talked about for breast, though, and colon. So let's make sure that we are comfortable with that.

Are both options, to the best of our knowledge, appropriate ways to read either chest X-rays or chest CTs? It's like in a deposition. I'm seeing nodding of heads, but if you could push the button and say yes, it would help a great deal -- or no.

I am going to take that as a yes from everybody that we believe it is concurrent. I don't know what else to do with it.

DR. GARRA: I would just like to comment. Since there is data out there that people have done it, I guess we have to say that it is valid. You know, it's like when you spot an abnormality as a radiologist now. We are just going to have to turn that part of our thing off and look for other abnormalities to avoid the problems.

CHAIRMAN GLASSMAN: Okay, any other comments? Yes?

DR. LEITCH: I would just say, you know, the

concurrent, you could say, would be more appropriate in the circumstance that you are looking for nodules. That is the pathology that you think is at hand; whereas, if you are looking for congestive heart failure, you don't turn the thing on first. You could do it after you have looked for those things if you are insisting on doing that. But the context in which you would consider doing it concurrently would be when your target is nodules.

CHAIRMAN GLASSMAN: Any other comments? Okay.

3. Do we believe that the users understand that the labeling is what the labeling is? Do we have a different opinion than we have had previously?

Let me remind you, previously we believed that if they knew what the labeling was, they often ignore it. It is like the old line somebody said about Richard Nixon when he was in the House of Representatives in the 1950s that if he had ever read the Constitution, he clearly didn't understand it.

DR. STEIER: Yes.

DR. BERRY: Yes, and we have validation of that.

DR. STEIER: Since there are a quoted study, Kobayashi and others, showing that reader detection -- it could be done as a second reader or it can be done as a concurrent reader -- I personally think that if that is out there, it should be applied for and labeled that way, and the labeling should be consistent with what is proven scientifically, and then people should comply with what is labeled as opposed to just going off and doing the way they think it should be done and not really following the labeling.

So I would be more comfortable encouraging companies to pursue proper labeling based on available science so that people could comfortably work within the labeling as it is published.

CHAIRMAN GLASSMAN: Are there any other comments? Let me then summarize -- Oh yes, Dr. Rosenberg?

DR. ROSENBERG: I would include training with labeling.

CHAIRMAN GLASSMAN: Let me try to summarize. The Panel believes that CAD is used for lung being both chest and CT, both sequentially and concurrently, that

both options based on the available science are reasonable and that we believe that if people know what the label is, they may not be following it and that the issue of again further training of users may help in that regard. Is that a satisfactory answer to this question?

DR. STEIER: Just one other thing, and that is that proper labeling should be pursued to match the current scientific data.

CHAIRMAN GLASSMAN: And that has also been mentioned. Yes.

MS. BROGDON: That's fine, thank you.

CHAIRMAN GLASSMAN: Thank you very much.

L7: Chest X-ray and chest CT are done for many important reasons other than looking for lung nodules. Can the use of CAD affect the diagnosis for these other conditions?

Can the presence of other conditions alter the effectiveness of the CAD function or the risk-benefit profile of the lung CAD device?

If the answer to either of these questions is yes, then are there specific conditions that should be represented by patients in the test database?