fatty or an age below which it is entirely dense. There's a lot of overlap and, if you were to hang 100 mammograms of women between 30 and 45 years of age, randomly distribute them on the viewer, neither I nor any other radiologist would be able to tell you how old the patient was from the density on the mammogram.

There could be many women in their thirties that are almost totally fatty replaced, and there could be women at 45 that have snowstorm dense mammograms.

I think the concept of enriching the population with women in their 40-45 age group is pretty standard in order to make the length of the study and the cost of the study reasonable. I think it is also true that most studies previously done have grouped all women under 50, pre-menopausal women, together, and I don't think there is any radiologic reason to expect that there is any differences between mammographic performance from 40-45 and in the thirties.

This slide was previously shown, but I just want to point out that the relative risk, assuming a

sensitivity of 70 percent rather than the 40 percent chosen by the FDA, yields a relative risk greater than that of ADH, atypical ductal hyperplasia.

I would also like to say that I think the FDA was perhaps pessimistic on the sensitivity of mammography in this age group. The DMIST study is an important study. It does show that in young women, women under 50, pre-menopausal women, and women with dense breasts, full field digital mammography performs better than film screen mammography.

So, certainly, I believe this group would be better served by that, but even the author admits that the 455 day follow-up was really an unusual, unconventional method of doing it. Also, in that study they used a 365-day follow-up to make the study more comparable to several other studies that are referenced within the DMIST digital mammogram study.

I think somewhere between 70 and 80 percent would be a more reasonable estimate. This is important because the post-test probability or the relative risk does depend on the mammographic sensitivity.

I think also that it would be wrong to assume that mammography is the only weapon in our armamentarium for evaluating the breast. Clearly, we are very effective with diagnostic breast Ultrasound, which is my area of expertise.

I can't say that I can advocate bilateral whole breast Ultrasound yet, but the ACRIN 6666 whole breast screening ultrasound trial is closed to accrual and will be completed within 18 months, and I am quite positive that those results will confirm the use of screening ultrasound.

There are future developments that are in the works. At least three combined fused, full field digital mammogram Ultrasound machines are in development. These are very exciting, because they offer the chance of getting both a whole breast Ultrasound and a full field digital mammogram in one visit, one room, one tech. It is economically compelling, and it is really about the only add-on test that actually offers the possibility of reducing callbacks, because many densities on the mammogram will be shown to be simple cysts.

Finally, there is MRI, which is really critical in the staging process, together with second look Ultrasound and mapping biopsies. But it can also be used in high risk patients, and we have already established the precedent of using this in patients who have ADH, LCIS, ALH, and this is approved by CMS and paid for by the local Medicare carriers, has no radiation risk, and it is equally effective in dense or fatty breasts. So it is not like mammography is the only tool within our armamentarium.

I think, in terms of risk, I think that you are not talking about a lifetime of risk here. You are talking about a single workup in the thirties, and that is not unprecedented. We already do that in women who have histological or family or genetic markers for high risk.

I think there is not a lifetime of risk as well, because they are already going to start mammography at the age of 40. So you are not talking about changing anything beyond the age of 40.

I think that we have other imaging methods in addition to mammography that offer no risk, and of

course, the benefit is new cancers detected at an earlier stage, and direct risks are pretty nil.

So in conclusion, I think T-Scan is a safe and effective technology, as long as you realize that it is a screening test and not a diagnostic test, and it meets an un met need in this 30-39-year-old age group; and I think that we have very accepted means for working up such patients.

I would like to make one comment, that in our population from year to year the percentage of all breast cancers that are diagnosed from 30-39 varies, but it is consistently double digits and rising. It is not four percent. Thank you.

CHAIRMAN CEDARS: And just prior to the last speaker beginning, I just wanted to remind you, you have about 15 minutes for wrap-up.

DR. GINOR: Thank you. Thirty seconds to correct the mike.

I would like to present you for a few moments with closing remarks.

We understand that this is quite a bit of information to digest from quite a number of experts.

We hope we have done a good job in presenting it to you.

I am aware of what was said earlier this morning about the small numbers, but there is a reason why we still rely on clinical breast exam. Without enriching, without developing methods to understand what we can offer patients safely and effectively other than doing 300,000 patient, 10-year studies, we will continue to rely on clinical breast exam and continue to live in an environment where physicians have to tell patients what surgeons often describe to me as the one-two punch: Yes, I know you are young, and you didn't expect to have breast cancer, and no, it's not early. You can't have a lumpectomy; you need a mastectomy; you need chemotherapy; you need radiation.

We are trying to avoid that in at least some number of patients, and with that regard 87 cancers in women under 45, 2000 approximately specificity exams in women under 40, is indeed a very large study and offers a very reasonable assumption of safety and efficacy with strong confidence intervals.

It is very important for me to explain to you that the primary endpoint was developed over almost a year in extensive discussions with experts and FDA, and that endpoint was exceeded virtually any way one looks at this study without stripping apart every single group into tiny little subgroups.

The benefit to risk ratio here is impressive. I don't think that it is my position, especially with such experts in the room, to undertake an extensive risk/benefit scenario, but the ultimate numbers are that we believe we will detect between 300 and 5000 additional cancers.

At the very worst, there are associations between, year after year, 1 million mammograms creating a potential 14 additional deaths. So the numbers there are very strongly in favor of additional screening, especially with modern mammography equipment and the nearly tenth dose of radiation as opposed to what was used years and years ago.

We do not see a product safety concern. The FDA did not either. This was a non-significant risk study based on a device that was already

recognized to be safe.

The additional mammography risk is minimal, negligible. Some of you probably recommend baselines to patients without any known risk factors at all.

There is a long-term, five-year study designed to address more thoroughly some of the questions that will be raised today and were raised earlier this morning, but there is no question that we have crossed the hurdle already, both in terms of safety and reasonable efficacy.

There are a few things that are going to be discussed later today, and I want to make sure that I address them for you and try to simplify them for you inasmuch as possible.

One of the issues that's been discussed with FDA extensively is the issue of algorithm development, algorithm stability, and that will be discussed with you extensively today.

I would like to draw your attention to two very, very critical components of this discussion. One, the questions of algorithm development do not

deal with the results in the pivotal study, but rather with what tests we could have, should have, maybe would have been prudent to have done before we started the clinical study.

The questions of algorithm stability initially deal with whether the company had enough information to start a clinical study, given the way that the algorithm was tested. We spent nearly 10 years researching this technology and developing this algorithm prior to starting the study, and felt that we had a reasonably safe algorithm, and we did, and I'd like to show you how we know that.

Initially, the algorithm was developed. Roughly 18 cancers were used, and the mean sensitivity that was found in testing that algorithm was approximately 34 percent. As you can see, the confidence intervals were wide, because that is a very small number of cancers around which to develop an algorithm.

Some of the CAD experts will tell you today that over the years very complex and better ways have been developed with which to evaluate an

algorithm prior to starting a clinical study. In our case, because the device was recognized as safe beforehand, we had the opportunity to test it in the market as opposed to test it on the bench.

The specificity originally was at a mean of 90 percent specificity with a confidence interval that was narrower, because we had more cases.

We then did a validation group where we took 12 additional cancers, tested them. The sensitivity was 29 percent with, as you see again, a very wide confidence interval. The specificity roughly stays the same.

We then started the PMA and handed the FDA an amendment on the first 90 cases that were collected in the Intent to Treat group. So some of these patients were excluded from the protocol analysis, and the sensitivity, as we expected, was again higher than 20 percent, and the specificity at that point has already stabilized; because by then we had 1,933 patients.

In the per protocol analysis we had 70 cancers. The sensitivity was 31 percent, again

confidence intervals 20-42. Specificity, as I said, had stabilized. And now we can offer you finally an analysis of what the sensitivity was on 87 cancers, the final datasets supplied to FDA.

This was bootstrapped hundreds and hundreds of times in order to test the stability of that result, and again the point estimate was 26.9 percent, sensitivity between 18 and 36 percent. That's two standard deviations, and 94.7 is the specificity.

So we are very, very comfortable at this point that the algorithm is stable. The FDA raises a point, which is not a bad point, which as we collect more data we will have to ensure that the algorithm becomes more and not less stable; and we have mechanisms in place to do that, and we will work with FDA as we collect more cases from the clinical world and build stronger algorithms to ensure that those algorithms are indeed stable and offer the expected results.

There are a few questions that you will be asked to deliberate about today, and I wanted to see

if I could offer you some insight into those.

One: The subgroup analyses that were carried out do not alter the conclusions. The threshold is exceeded well, no matter what subgroup analyses are performed.

The other thing that mentions is that the level of risk that is identified by T-Scan patients significantly exceeds the standard of care at which we currently offer additional screening, and that does not change, irrespective of the analyses performed.

There are some questions now retrospectively about the enrichment which was agreed to at the beginning, and we wanted to explain again. I think most of you will agree, age in and of itself -- that is, including women 40-45 as opposed to women just 30-39 -- should not have an impact on the breast tissue, assuming they are all pre-menopausal, that EIS technology is independent of age. There was not a significant difference in sensitivity between younger and older patients.

It is very important for me to express to you that we are aid to clinical breast exam, and our

main focus is on the sensitivity in smaller lesions. The sensitivity in smaller lesions is better, not significantly but much better, than in larger lesions.

The same is true for palpability. And again, one of the questions you will be asked is to evaluate Israel versus U.S. separately. We don't believe that there is a statistical reason to do so. However, we do understand that there is a legitimate concern about how this is going to be utilized in the American patients.

Again, Israeli patients were Caucasian, which represent the largest group in the American study. There is absolutely no reason to assume that the results will be different between Israel and America, and again if Israel is looked at independently and America is looked at independently, still all the success thresholds are beaten.

The questions that are posed by FDA have to do with technical difficulties, and I really do want to explain this to you, because it is very, very important.

The FDA had a suggestion, which was a very

smart suggestion, at the beginning of the study, that we should blind those sites that are going to be performing biopsies on patients. The concern, which was a legitimate one, was that if physicians use a technology as part of a clinical study, they generally believe it is a relevant technology, and we were afraid that if a lesion didn't look very suspicious and a T-Scan said that it wasn't very suspicious, the physicians might not carry out as aggressive a biopsy or as aggressive a treatment program as they otherwise would.

So they were entirely blinded to the feedback that one gets when using the device in clinical practice. So when you see the difference in the fact that we had virtually no technical difficulties in the specificity arm where physicians were not blinded and almost 10 percent or nearly 10 percent technical failures, which was one site, in the sensitivity arm, that is something that could not happen in clinical practice.

It happened, because the physicians using the device didn't have feedback to tell them that the

device wasn't working appropriately and, therefore, were unable to call us and say could you send a technologist over, there appears to be something wrong with the device. This is not something that can happen in "the real world."

There are some discussions about whether the existence of T-Scan could have an impact on risk in regard to mammography. I would like to highlight for you again, we are not here today to debate the standard of care in general. We are not here to debate whether mammography is or isn't a good tool.

We are here to debate whether or not T-Scan crosses the threshold at which we offer screening to other women, and it does so by nearly a 300 percent margin.

We are currently offering in America screening to women 40-49 who are at absolute risk of approximately 0.0029. The T-Scan's absolute risk is about 0.0073. Therefore, we believe that this is not a discussion about benefit to risk. It is a discussion about whether or not America wants to change the standard of care. Until it does so, this

device appears to make clinical sense.

Again, mammographic screening of T-Scan positive women is more than three times as effective as the current standard of care, and we hope that that message will carry on for the rest of the day.

This is the conclusion of our morning presentation. We appreciate the time and the patience that you gave us to share this data with you.

CHAIRMAN CEDARS: Thank you. I would like to open the discussion now to questions the Panel might have. Dr. Glassman?

DR. GLASSMAN: Yes. I'm not sure who to pose this to, maybe Dr. Stavros or one of the other experts.

As a clinician, screening is a process. It starts with a test, and it ends with a diagnosis, if you can make one. It is not just a positive T-Scan. So I actually have several questions.

One: If there is a positive T-Scan in a 30-year-old woman or a 35-year-old woman, does she ever get another T-Scan or is she positive for life?

Second: If the patient has a positive T-

Scan, goes on and gets a mammogram, has a dense breast, goes on and gets an Ultrasound, nothing is found, then she may get an MRI. Is she, in your expert opinion, going to be an annual screening patient in her thirties because of this positive result, because it goes to the issue of cost/benefit.

DR. GINOR: Thank you, Dr. Glassman. If you don't mind, what we did yesterday is extensively debrief our experts in regard to questions we thought were coming and, therefore, I wouldn't mind answering your. But if you would like a more thorough answer afterwards, I'd be more than happy to ask Dr. Stavros to come up.

What we have discussed and what we feel is because we measure a physiological exam at a particular time, we believe that this is a one-time risk as opposed to a lifetime risk. There could be something going on in the breast which is perhaps not neoplastic, and we do not want to ascribe risk for life to a particular patient.

Therefore, we believe that, if the way that the radiologist, should that be where the patient

is sent, performed an Ultrasound or a mammogram and those are normal, that patient would get a T-Scan again as part of her annual exam next year. We do not feel that we yet have multi-year data to suggest that that patient should be treated differently for the rest of her natural life.

The same is true, no matter how far you go. For example, in the U.S. Army originally patients only received mammography. Over time, the radiologists became very comfortable and confident in the T-Scan result, and they now have decided to perform MRI.

We are not in a position where we can legally debate or discuss or suggest what the follow-up community is going to do, but we do have the responsibility to assume what that might be, and we believe it is probably going to be digital and/or Ultrasound, and ultimately MRI, if the results are as good as we believe they are. Does that --

DR. GLASSMAN: It answers my question, yes.

DR. GINOR: Thank you, sir.

CHAIRMAN CEDARS: Dr. Mortimer?

DR. MORTIMER: I'm curious. Could you just define -- maybe I missed this, but what a normal breast exam was, is one question.

Secondly, are there differences in impedance that correlate more likely with a positive finding? I realize again the sample size is fairly small.

Thirdly, in your histologies, there was no lobular carcinoma in situ identified, and I realize again these are small samples, but I also know that you have follow-up data in the previous scan results.

DR. GINOR: So there are three questions, I believe. One is what do we consider a normal clinical breast exam? What we mean by that is, if the physician in performing his or her clinical breast exam on the patient finds an abnormality, that patient is now at a level of risk where they deserve additional follow-up, irrespective of T-Scan.

Therefore, those patients have a normal CBE, and then generally would be sent home. Is that different than the question you are asking?

DR. MORTIMER: Well, so were fibrocystic changes considered to be within the normal variant or were they considered an abnormal breast exam?

DR. GINOR: Unless there was a report of an abnormal clinical breast exam, lump, mass, nipple discharge, etcetera, if that patient would not have been sent forward for additional workup if T-Scan wasn't there, that was considered a normal exam.

We did find one LCIS, and so I really don't think that we are in a position to be able to build that out, and we also did not consider that, and malignancy -- our agreement with FDA is only true malignancies would be considered malignancies in the study, and LCIS is sort of on the fringe, as you know.

The majority of what we found -- I would say 82 percent of the cases we found -- were infiltrative ductile carcinomas.

Your second question I wasn't 100 percent clear on. So I apologize. I'd like you to repeat it, please.

DR. MORTIMER: Are there variations in the impedance measure that correlate more likely with

cancer finding?

DR. GINOR: Yes. I mean, that is the basis of the technology. What one assumes is that areas of change, especially areas of change that incorporate changes in the amount of cellular fluid in and outside of cells, that is exactly what the T-Scan finds.

Those graphs I tried to show in the beginning perhaps were not exhaustive enough, but it is that delta between normal tissue and abnormal tissue that we recognize as atypical; and if it is recognized as atypical, we believe that is a level of risk, as was shown in the study, that warrants further workup.

CHAIRMAN CEDARS: Dr. Hillard?

DR. HILLARD: I heard agreement among the speakers that menopausal women should be excluded from the studies, and were. Can you tell me the definition of menopause?

DR. GINOR: Wow.

DR. HILLARD: What was the definition for the study? How was it defined, and what then is the

impact of including or excluding women who were or were not menopausal?

DR. DICKERSON: Let me try to take that. A study definition was six months of amenorrhea. Yes, sir. I was on the panel, remember.

The study actually used a definition which we don't use in practice. In general, obstetrician gynecologists use 12 months of amenorrhea to define menopause. In this particular case, we actually used a more stringent criteria of six months. If there were six months of amenorrhea, then the patient was considered to be menopausal and was excluded. Even though that probably excluded some that we would call perimenopausal, it was slightly more stringent than otherwise.

There were not laboratory evaluations, which as you know, depending on the day, the time and the place in a possible cycle, might vary in a perimenopausal woman.

DR. HILLARD: So that a woman who was on extended cycle oral contraceptives and had not had bleeding for six months would have been excluded or

included?

DR. DICKERSON: Well, you are luckier with those than I am. I -- go ahead.

CHAIRMAN CEDARS: Can you come to the mike, please?

DR. DICKERSON: Six months without a period and not on hormones, they were excluded.

CHAIRMAN CEDARS: And can you introduce yourself as well?

DR. LENINGTON: Certainly. I'm Sarah Lenington. I am the Director of Clinical Development for Mirabel.

So it was women who weren't on -- having hormonal suppression of their periods, who had not had a period for six months that were excluded.

CHAIRMAN CEDARS: Dr. Snyder.

DR. SNYDER: In reviewing all the material, I never did see anything related to the reproducibility of the data in a single patient. So I'm curious, you know, if you have a green light, you know, and you repeat the scan 20 times, what's the chance of getting a red light in there, and vice

versa. If you have a red light, what is the chance you are going to --

DR. GINOR: That is a very good question, Dr. Snyder. In fact, that question was raised by FDA as part of their initial review, and we were asked to conduct what was termed a repeatability study.

We conducted, I believe -- Let me just make sure whether it was 100 or 90 in the repeatability study. We took 10 women and put them through 30 exams and measured repeatability. That was handed in to FDA, and there was no statistical difference at all in repeatability.

What we did was different machines, different examiners, different women rotating through. So we had three devices. Ten women rotate through over and over again to see whether there were changes, and there were not.

DR. GLASSMAN: In that repeatability study, were some of the women positive and some negative or all negative?

DR. GINOR: One woman was positive, and she was positive repeatedly.

MS. MAYER: Did you study variability or, rather, stability of electrical impedance over a woman's menstrual cycle?

DR. GINOR: That's a good question. That question was asked, actually, as part of our initial approval back in 1999, and we conducted a study evaluating that. Indeed, there are differences in electrical impedance, depending on time in cycle.

When we had the initial device, which was a very high sensitivity device, in fact, that device was sensitive enough to identify those changes. That element is muted in the current device, which is why sensitivity is lower, because we raised the threshold for what is, frankly, considered a problem to the point where those variabilities in menstrual cycle no longer matter enough to make the difference from negative to positive.

What I mean by that, if I can be a little more specific, is in the old device there was an image, and changes in menstrual cycle could be almost identified on that image. You could wonder why that breast looked different.

That is no longer the case here where everything is measured electronically, and either the measures are below the threshold or above, and that threshold is too far from baseline to be pushed over just by menstrual changes.

MS. MAYER: So just to follow up, did you find that you had less sensitivity with the raised threshold?

DR. GINOR: Not on this device.

DR. TAUBE: So in terms of reproducibility, why -- How do you explain the fact that a single site accounted for about 93 percent of the failures in your technical difficulties?

DR. GINOR: A single site had a broken device. Actually, a single site had two broken devices. In regular practice, you would do an exam, and you would recognize -- There would be this message saying insufficient data, and the result would come up, insufficient; and you would call us, and you would say something isn't right with the device.

We used a piece of software -- We wrote a piece of code right before we initiated this study so

that you would not see that in the sensitivity sites, and so that one site had two broken devices in shipment, and we didn't know that that happened until we went for our routine monitoring there months later and saw that they had never had a positive, even a false positive exam, where on average in those women there are a significant amount of false positives, because they are waiting for biopsy.

That's what triggered to the fact that something was wrong with that device. But in clinical practice, you would know within one or two exams that something wasn't right. You would call us. In fact, in the U.S. Army study this happened several times where somebody dropped a probe, and immediately they recognized there was a problem and called us.

So that's really not a matter of variability. That's a problem of the fact that there was a monitoring window between when they got the device and when we caught onto the fact that there was something wrong.

DR. TAUBE: So that site wasn't involved in the specificity study?

DR. GINOR: Right. It was a pure sensitivity site, and it was a site that we added to try to have more cases in America. We had nearly twice as many sites in America than in Israel to try to have more American cases, but we found it very, very difficult for patients in America on their way to biopsy, you know, an hour before, to partake in a clinical study, and we tried very hard not to make those women feel any pressure and, therefore, we didn't really have that many cases.

DR. TAUBE: What was your denominator? I mean, you said lots and lots of women didn't want to participate. How many women did you actually ask?

DR. GINOR: It's not really -- I should have been a little bit more careful in stating that. It's not that lots and lots of women didn't want to participate. It's that in the United States the way they were scheduled for biopsy didn't always give us enough time.

In America, I believe -- and I think this is probably a good process -- many of the centers are now trying to take women, diagnosis, treatment, and

give them a full answer within one day; whereas, in Israel there was a lot more time between when the patient was scheduled for biopsy and when the biopsy took place, and we had a much bigger window in order to ask women to participate.

Here, what we found, that some of the centers were -- that patients were just moving along very quickly. The physicians were very, very concerned about getting them to their biopsy, getting them out of their biopsy, and we just didn't want to interfere in that pathway. It was something that we really were careful about.

DR. TAUBE: Can you give us an idea of what the denominator was?

DR. GINOR: Can we answer that question after the break, and we will pull it up and see?

DR. TAUBE: Yes.

CHAIRMAN CEDARS: I do want to remind the sponsor that they will have time after lunch, if they need to look for some of this data.

If we could move on, we just have about five more minutes for questions. We will go around

the table this way. Dr. Snyder.

DR. SNYDER: You know, now I've heard three explanations for the technical difficulties at this one site, and the reason I'm going to ask this is because it -- In the specificity study, there was one percent technical failure. I mean, that is probably what we are looking at.

So I understand the deviation of like 12 percent in the sensitivity study. But from a potential user of the device, why didn't they get an error message? In other words, you know -- and I understand the process by which a machine works, you know, that you are told whether you've got adequate pressure on there, and there appear to be a number of different steps where you are told whether you are performing it correctly or not.

So why didn't at that one site they get an error message?

DR. GINOR: Okay. I'm going to try to explain this more specifically, because it is very important that this is properly understood. If I haven't done it right yet, I'll keep trying.

In the specificity arm, we were not going to change patient management, no matter what the result was. So there was no "risk" in showing you whether she was red or green, because we were not going to change her management based on a device that is still under evaluation.

There was a concern that we could potentially bias a surgeon prior to a biopsy by a negative result. So all visual feedback was eliminated. So that did not give you the opportunity to recognize that you had a problem.

Therefore, up until our monitoring visit -- I think it was two or three months later -- came in and said something is wrong with this device, we had no indication. In clinical practice, nobody is going to ask us for a device that is blinded. So they are going to know within one exam that something there isn't operating right, kind of like when something is wrong with a transducer on Ultrasound or something of that nature.

CHAIRMAN CEDARS: Dr. Miller.

DR. MILLER: I have a couple of questions

that deal with things that really haven't been raised to this point.

The first is: Is there -- I didn't see any sub-analysis of the range of grade of tumor and stage of tumor in terms of how that impacted the sensitivity and specificity, and specifically in the enrichment group. Is there any a priori reason to think that the type of tumors that would be suspected in the 30-39 group would be different than the 40-45-year-old group, since that enriched group was included?

The second question is: It seemed to me like there was a difference between the cup size for the Israeli population versus the U.S. population. I would like some further understanding of, if this technology were deployed throughout the country, is it, in fact, the case that this technology is much more sensitive in smaller breasted women than it is in larger breasted women, and would that need to be a disclaimer? Would that need to somehow modify labeling?

The third question has to do with: Is

there enough ethnic diversity in your population? Since you don't really speak to the ethnicity of the Israeli population, but you have said here that it was primarily Caucasian, the subsets of ethnic groups that are disclosed are relatively small by comparison, and that could have, again, impact on breast size, sensitivity and specificity, and may have -- Again, this is not my area of expertise, but may have an impact also on tumor types.

Finally, is there enough information about the hormonal milieu of women who probably are still dealing with contraception between 30 and 39 as it relates to impedance and skin effects? Actually, I'm interested in a broader part of this, which is this is also a time in life where women are concerned about dermatologic conditions. What other factors influence impedance, skin impedance and tissue impedance, that are germane to our understanding of the limitations of this technology even as a screening device?

CHAIRMAN CEDARS: And if I could just request, because we are running short on time -- I saw you writing those down -- if we can at least make sure

we have all of the questions out in the open so that you and FDA will have a chance to answer them, I would appreciate that. So if you could hold on your response, if that's okay, Dr. Miller.

DR. MILLER: Yes, that's fine.

CHAIRMAN CEDARS: Dr. Berry?

DR. BERRY: So on this 1 in 666 or 1.5 in 1,000 figure, this is, as I understand your presentation -- this is the prevalence in the decade. So it is the proportion of women, cumulative proportion, who will have breast cancer over their thirties. Is that correct? I think it's correct.

DR. GINOR: It's not cumulative. The cumulative effect is the data that I reported from NCI, which is that 229 -- one in 229, which is nearly four per thousand.

The prevalency that we are talking about -- and that's the dichotomy between SEER and what we are saying -- is that the studies we took are studies that took, for the sake of discussion, 1000 women and offered them mammography under a clinical study, women who didn't necessarily have family history, but it was

a way for the scientists who undertook these studies, some of them very large studies, to see how much cancer was really there; because things like SEER don't answer that question, because we don't know how much --

DR. BERRY: So what is the 666 figure?

DR. GINOR: 1.5 cancers per 1000.

DR. BERRY: Per what? Per year for any age, for age 30?

DR. VERTER: I believe -- Joel Verter. I believe the correct interpretation is what epidemiology would call a point prevalence, that at any point in time, if you screened all the women, that is the expected number of cancer cases you would expect.

DR. BERRY: But, Joel, the rate increases by tenfold over the decade.

DR. VERTER: Sure. So you can think of it as an average over the decade, if you like, but --

DR. BERRY: Well, my figures, SEER figures, are a good deal less than that.

DR. VERTER: But SEER, I think, is

incidence, and we are talking about point prevalence. So those are two different concepts, right, incidence and prevalence?

DR. BERRY: Yes.

DR. VERTER: Okay.

DR. BERRY: But in the figures that you are using -- Okay. I think the rates are a good deal less than that. That's about the total over the decade. But let's ask what I think is a more important question.

Something that I don't understand is the blinding in the study. I heard some statement about blinding. I am concerned about, in particular, the sensitivity study. The operators -- and you tell me that it turns red or green, but it surely has some operator effect. You can linger longer on a part of the breast. You could go over it again.

Who was blinded? Did the operators know that these are women who are going for biopsy?

DR. GINOR: I'm glad you brought that up. I think it is one of the areas that we are most comfortable with in this study, and I will tell you

why. First of all, yes, every single woman in the sensitivity arm was known to be going for biopsy.

Secondly, there is no way that an operator under any circumstance, including myself who knows the in and out of the technology, can force the device to go red or green. Furthermore, the FDA put a control in place, which was a very smart control.

They wanted to see what our false positive rate was in the sensitivity arm, and let me tell you why that is so important. Let's say that you are really enthusiastic about this technology. In fact, you are my brother-in-law, and you really, really wanted to ensure that all cancers were found.

You would, theoretically, if you could manipulate the result, have a lot of reds. Of course, the specificity would be very, very low, because you have no way of predicting in advance which cancers -- which biopsies are going to be malignant and which are not. The hit rate, by and large, on biopsy is about two out of 10.

So first of all, it's impossible, purely impossible to change the result. You can't linger

longer. You hit Record. The device collects the data, and it analyzes it, and you don't know if it came up red or green. So you can't say, well, gee, I sure think she's positive; let me go ahead and re-record this. And you are doing this before the biopsy. So you have no way of saying, gee, I'm certain this is a cancer, let me record again and again.

Now let's say you've finished the exam, and for some reason or another you felt that you wanted to ensure that it came up red. We would then have a second case in the system, and we were audited thoroughly, and we didn't have additional cases over and over in the same patients.

So there actually is three kinds of blinding. One, you didn't know what the result was. So you didn't know if it was red or green and whether you should manipulate it; (b), you can't manipulate it, and (c) you do it before the biopsy.

DR. BERRY: Have you looked at the operator effect in a reproducibility study across operators?

DR. GINOR: Yes.

DR. BERRY: And what do you find?

DR. GINOR: There was no significant operator effect, and --

DR. BERRY: There was no operator effect or there was no significant operator effect?

DR. GINOR: There was no significant operator effect.

DR. BERRY: So there was some. Sometimes an operator could do the same woman and get a different result than another operator?

DR. LENINGTON: The study we did looked at the effects of both devices and operators and found no effect, in fact, for either devices or operators. The P-value -- I don't recall exactly what it was, but it was very, very high on both of those issues.

DR. BERRY: I don't know what the P-value would refer to, but there were some cases where the operators differed?

DR. LENINGTON: No. There were some cases that turned out red, and it wasn't that there were cases where operators differed. There were some cases

-- individual cases that turned out red, but those were randomly distributed across operators.

DR. BERRY: I don't mean that. You do two operators, same woman. One gets red; one gets green. Did it ever happen?

DR. LENINGTON: Occasionally, that happened, but then when the cases were repeated, maybe the other operator would get red and the other one would get green. So that there was a random distribution of red cases across operators.

DR. BERRY: Okay. So finally I want to ask about the -- The FDA presented data that said that in the U.S., of 13 women that had breast cancer, your device detected none of them. This is in the age group 30-39.

So what you are saying is that you are asking for an approval to detect cancers for women in their thirties in the United States, and you have never detected one.

CHAIRMAN CEDARS: And again, just so we can get all the questions in, if we can hold that question. Dr. Weeks?

DR. WEEKS: Yes. Could you please -- My question is: Why does the sensitivity of the device seem to be better with smaller lesions?

The second question is: Does the device perform differently -- We saw data on cup size, but it seems to me that the electrical current is being transmitted through tissue, in some cases, other than just breast tissue. So does body mass index affect the performance of the device, and was there a significant difference in body mass index in the U.S. subjects versus Israeli subjects?

CHAIRMAN CEDARS: Let me just -- Again, if we can just get all the questions through, some of this may be addressed in the FDA response, and then the sponsor will also have time for additional questions.

DR. BAILEY: I will just add this. What we are going try and do is try to ask you a number of questions now. You can write them down. I think we probably would like you to go ahead and respond to Dr. Berry's last response, but what I would like our people to do, so we can keep the process moving this

morning, is for our people, if they have any questions they would like answered, have the company write them down and right after lunch we will give the company a chance to respond to all the outstanding questions we have not gotten to.

So first I think we are going to let them respond to Dr. Berry's last response, because I think that they already -- and then we will start going through and sending the rest of the questions around, and then we will move on.

DR. GINOR: I apologize. I didn't recognize it was a question. I thought it was a statement, and if I had recognized, I would have answered.

We did discover cancers in the United States. When the subgroups are ripped apart, although that was not the way the study was originally designed, and you eliminate patients that were supposed to be included by the study design, you can make that argument.

I remind you, this was designed to be a multi-center international study without limitations

on whether the data would come from Israel or America or anywhere else, and the fact that there is an ocean between Israel and America shouldn't really make a difference unless one thinks that the physicians there are less good or the patients there are different or the practice of medicine is different. But this is a situation where (a) none of the clinical features were different.

The practice of medicine is not different in this regard, because a biopsy is a biopsy. The pathological report is a pathological report, and the physicians were trained by the same team, and they were of similar experiences.

So I'm not really certain that that is a fair comment to make, and I'm not really certain that that is the question, but in fact, the study is a uniform study across both countries.

CHAIRMAN CEDARS: Okay. Additional questions?

DR. GOLDBERG: A couple of questions here. On one of the earlier slides you talked about 2000 women were screened. 1900 were normal, and 100 were

T-Scan positive, and that subset indicated a five times greater risk of breast cancer.

The question was: What percentage of these actually went on for interventional workup and were confirmed to be positive for cancer?

The second question, not to beat the issue too much, was back on the exclusions. There was a one percent number. There was a 12 percent number, and one of the slides said that 55 patients were excluded for technical difficulty. Just to clarify if that was a probe or mechanical failure or was that a technical difficulty because of the patient.

Third question was: In the enriched group of women, the 40-45 age group, were these women also additionally screened with routine mammography?

Then the fourth and last question is regarding a mechanical question. In the data there were nine areas of the breast that were scanned. Did the nine areas vary with a variation in breast size? Was the probe able to make adequate contact with the breast on small breasts, and were there overlapping areas of the nine areas in each breast, again

depending on breast size? Thank you.

CHAIRMAN CEDARS: Okay, and I believe we had some additional questions here.

DR. ROMERO: Yes. I think I'd like to follow up with some of the questions that Dr. Miller asked, particularly with regard to ethnic/racial subgroups, but on a slightly different tack. So if you could just, for the purposes of responding to my question later, I am thinking more in terms of --

CHAIRMAN CEDARS: Could you come a little closer to the mike, please.

DR. ROMERO: -- more in terms of recruitment into the study and the methodology that was employed. It seems that the racial/ethnic distribution of both samples in the sensitivity and in the specificity arms were really not representative of the general population, and it is quite discouraging to see that racial and ethnic minorities who have not been included in relation to or in proportion to their representation in the country in the past, as has been acknowledged by a large study supported by the Federal government as well as conducted by individual

organizations, has sought to remedy that problem.

So I find it quite discouraging to be looking at data that sort of repeats some of the errors of the past, and it makes it difficult to interpret your subgroup findings when you have done statistical analyses that either indicates, for instance, in the specificity arm that there were not significant differences -- No, actually, there were significant differences by race/ethnicity, although it is not dichotomous.

So it is impossible for us to see where, between which subgroups, there were differences. Then in the sensitivity arm where differences by race/ethnicity did not appear, but then again the numbers in those subgroups are really dramatically small.

So if you could pretty much discuss what, if any, efforts in the sampling were taken to try to have representative proportions, that would be helpful, and probably more important, what the implications of these sampling numbers are with regard to -- clinical implications with regard to the use of

this technology in women of different racial and ethnic background.

CHAIRMAN CEDARS: Two additional brief questions or comments.

MS. MAYER: A question again about the prevalence figures. Were those -- Was that 1.5 per 1000 prevalence based on a population from whom the clinical breast exam positive cases and family history cases had been removed? In other words, was it your true apparently low risk population? That's one question.

Then just to follow up on Dr. Romero's question, I wonder if you could address if, in fact, that 88 percent specificity figure for black and Hispanic women is something that you feel confident of, what the implications might be for that population in terms of false positives? Thanks.

CHAIRMAN CEDARS: Dr. Hillard.

DR. HILLARD: Just quickly in following up the previous questions about the anatomy of the breast and the geometry of the breast in terms of the nine areas that were sampled and differences in women's

breast size, if you can clarify instructions about assessment of the axillary tail of the breast and whether that differed in women with different breast sizes, as well as the sites of the biopsies and the masses that were found.

Clearly, with mammography those areas are areas that are not as well sampled. So any differences, and comment about that with this device. Thank you.

DR. TAUBE: Since the cost/benefit analysis depends on the assumption of benefit, have you looked at data to indicate that there might be greater success if you were to find tumors stage by stage in younger women that the outcome, the treatment and the intervention and so on, would be effective?

CHAIRMAN CEDARS: Okay. Just briefly, clarification here.

DR. BERRY: Just following on that and Dr. Miller's question, more important than stage is -- in terms of treatment and prediction is ER status. Would you tell us the estrogen receptor status of the cancers that were detected by the device?

CHAIRMAN CEDARS: Okay. We will now need to take a break. We will cut the break short to just five minutes. I also want to remind the Panel members that they should not at this break or any future break discuss the PMA amongst themselves or with the sponsors or with participants.

(Whereupon, the foregoing matter went off the record at 10:43 a.m. and went back on the record at 10:54 a.m.)

CHAIRMAN CEDARS: Again, if we can ask everyone to take their seats, please, and as people are taking their seats before the FDA begins its presentation on this PMA, if I could ask -- Nancy Brogdon wanted to speak.

MS. BROGDON: Yes, thank you. I just wanted to provide a reminder to the Panel. We heard the word cost a couple of times, and I think we understood what was meant was risk/benefit as opposed to cost/benefit. I just wanted to remind the Panel that the agency and the Panel cannot take into account economic considerations.

So if you mean risk/benefit, please try to

say risk/benefit. Thank you.

CHAIRMAN CEDARS: Okay. We are missing just a couple of people from the Panel. We will give them just a few more minutes. Here's Dr. Weeks.

Nancy, if I could just ask you to reiterate your comment, please.

MS. BROGDON: I just wanted to remind the Panel that the Panel and the agency are not allowed to take into account economic considerations. So if in your discussions here you talk about cost/benefit, we would like you to be very specific that you are talking about risk/benefit, if that is the case. Thank you.

CHAIRMAN CEDARS: Thank you, and with that the FDA will begin their presentation with Dr. Robert Phillips.

DR. PHILLIPS: Well, good morning, and thank you for coming to assist us in reviewing this particular device.

The topic again is the T-Scan 2000 ED, and I am Robert Phillips. I --

CHAIRMAN CEDARS: Could you bring the mike

a little bit closer, please.

DR. PHILLIPS: Sure. I am Robert Phillips. I am the Chief of the Radiological Devices Branch.

The indication for use you have seen before, but let's refer to it again. It is indicated as a complement to the Clinical Breast Exam in asymptomatic women age 30-39 (inclusive) with a negative CBE and a negative family history for breast cancer.

As you heard from the company and other discussants this morning, this device is looking to function in an area where there is very little else at the present time for you to use.

To refresh your memories, we put together a chart here of what are the current guidances for the use of -- for the standard of care for breast cancer for average risk women, and there's -- Well, four different groups are on this chart.

As you can see, for women who are less than 39, the recommendations are consistently CBE and nothing else, and then for imaging there is some

controversy about whether you do it every year or not at 40 and above, and whether you do it differently from 40-50 and 50 on, but that really is not applicable to this situation.

You are going to hear talks today from several of our reviewers. The device description and pre-clinical data will be presented by Dr. Kish Chakrabarti. A discussion of the algorithm stability will be given by Dr. Nicholas Petrick. A clinical review will be by Dr. Ron Yustein.

A statistical review will be by Lakshmi Vishnuvajjala, and the risk/benefit analysis will be by Dr. Roselie Bright.

With that, we had a few other reviewers who were involved with this, who are not going to be talking. They are Dr. William Sacks who assisted us in clinical; Dr. Harry Bushar in statistics; Dr. Robert Wagner in assessment of the algorithm; Joe Jorgens who did the software review; Kevin Hopson who is involved in our biomedical inspection program; and Fleadia Farrah who was involved in the good manufacturing processes.

With that, I will pass this on to Dr. Kish Chakrabarti who will talk to you about the device overview.

DR. CHAKRABARTI: Good morning. It's still morning. I am Kish Chakrabarti. I am a physicist, and lead reviewer for this submission.

I am going to talk about very briefly device description and operation. You already heard several times from the sponsors that this device -- It analyzes multi-frequency capacitance and conductivity when conductivity is inverse of resistivity, of which you had more than conductivity, and 8 by 8 sensors -- You saw the picture of that -- at 17 preset frequencies.

Then the results are based on the scan from both breasts. Device does not show or identify the location of any suspicious region in either breast.

The device provides a binary outcome, as you heard, of negative or positive. You also saw a picture the sponsor showed you. The solid green line indicates negative, and hatched red line indicates

positive. Device does not produce any image for diagnosis.

I will just simply say a few points on the previous device that was approved. It was called T-Scan 2000, and that was the PMA number for the approval. That device was approved as an adjunct to mammography with equivocal Bi-RADS assessment 3 or 4, not for cases with clear mammographic or non-mammographic indications for biopsy. It was not a screening device.

Target population was age 40 and over. Attending physicians determine if T-Scan should be used, and attending physicians also interpret results.

The devices uses a different frequency range. It uses a different algorithm and converts computed capacitance and conductance to gray scale.

The device produces an image, displays shades of gray, provides a bright region on displayed image for malignant tumors compared to its surroundings.

I am switching back to the proposed device. The bench studies: The device complies with

the International Electrotechnical Commission standards, IEC 60601-1 and IEC 60601-1-2. The test results and the reports that sponsors provided are satisfactory.

Pre-scan check: T-Scan performs pre-scan safety tests each time the device is turned on.

Software safety tests were also satisfactory and acceptable.

Biocompatibility and animal studies: These studies were performed in the previously approved device, and these are the same handheld surface and signal transmitter. So we did not need any further data on those.

Next is algorithm stability as a part of the preclinical data. Dr. Nick Petrick will present that.

DR. PETRICK: Okay. So I will continue on preclinical studies and, in particular, I am going to really concentrate on the algorithm stability analysis that the sponsor conducted. Just as an overview of my presentation, I will talk about the algorithm architecture and implementation. I will

talk about the training and validation datasets utilized by the sponsor in this analysis, and then talk about the stability analysis that was conducted.

First an important T-Scan architecture: It's made up of two components. It is made up of a hardware component or the probe section of this device, and a software algorithm. The algorithm is the recipe for making the decisions of suspicious or not.

It uses trained -- It is trained using patient data. So the algorithm only works, because of its training process. It is integral to -- and this software is integral -- or algorithm is integral to the device operation. So we have to keep in mind that this device has both a hardware and a software component to it, and both are important.

Just a block diagram. Here I show the probe system and just a block diagram of this. It produces impedance measurements, some number of impedance measurements. They are digitized into a digital form. Those are used as features or sets of features that are put into some sort of prediction

algorithm, and an output is produced. In this case, it is a binary suspicious or not suspicious output.

The implementation again: The patient is examined with the T-Scan probe. This information is then imported into the algorithm where a classifier is applied to determine a score for each breast. A threshold is applied to that score to determine suspicious or not for each breast, and then a logical OR operation is applied where, if either breast is suspicious, the patient is considered suspicious. Again, this goes from a hardware scan through to a binary decision on the patient.

The device uses impedance measurements. There are two impedance measurements, 17 frequencies, nine sectors, for a total of 306 impedance measures per breast. This is a fairly large number of initial features, which is one of the reasons that algorithm stability becomes an issue in this device, is with a large number of features a algorithm has a potential to be unstable, but we want to take a look at that and see if this algorithm had any issues.

The algorithm training process: Again, in

order for this algorithm to perform at all and in order to look at stability, we have to talk about the training process. In order to do training, you need a training dataset. The sponsors call this the learning group.

The next phase in developing the algorithm is to do dimensional reductions where they take those 306 measures to reduce the sum set of blended features, and this is a multi-step process.

Once there is a limited -- Once the dimensional reduction step has been performed, then weights are determined for the classifier. So it determines blended weight and feature weights for each of those features that are going to be used in the software algorithm.

Following this, there needs to be a threshold selection to select a cutoff value between suspicious or not, and this results in the trained algorithm where the algorithm is trained to appropriately score similar patients as in the training set. So this process is important in understanding how the algorithm is developed, and then

how it is evaluated.

So the next question is what is the stability or what is algorithm stability or what are we talking about in this measure. The definition that we are going to use is a measure of uncertainty in algorithm performance with variations in both the test and training data.

Then in particular I have underlined the training data here, because one of the issues -- The pivotal study will look at test variability, but what we are interested in is what happens if there were differences in the training set to see how stable the overall algorithm, the architecture of the algorithm, is.

The stability of the T-Scan is related to dimension reduction process, to the estimation of the algorithm weights, to the estimation of that cutoff threshold, and importantly, to the number and quality of the training cases used to develop the algorithm.

So I just want to give you some indication of why stability analysis may be important. One of the reasons it could be important is it indicates if

the stated performance is due to a fortuitous choice of training and test set.

So here I just showed an example. This is not data from the T-Scan device. It is just an example, a hypothetical example to try to give you an indication of why this might be important. We can look at a performance for some device, and it has some set of error bars associated with it. These are test confidence intervals maybe for this device.

We are also interested in what happens if we had a new set of training cases? How would that algorithm perform overall? We can look here. What I've just done is put on -- again, these are just an example -- a set of training error bars associated with that device. If these error bars are fairly well constrained, we would call that algorithm maybe more stable than some other algorithm.

If, on the other hand, those error bars are very, very large, which means that if we had a new set of data, we would have a very large possibility of performance estimates, we would call that algorithm less stable.

A second reason for wanting to understand algorithm stability is that algorithms evolve over time as more data is collected. Now it is important to note that this T-Scan device is a nonadaptive algorithm. It doesn't change as a new case comes in immediately, but what we see within the FDA is when software revisions come in, they come in because they have a -- sometimes with a new software and new algorithm that is produced.

So software revisions produce evolving performance estimates of performance for the devices. If we have a more stable algorithm -- So what I'm just showing on this plot is different time points. This would basically indicate new or different training, hopefully additional training cases that may be used in the algorithm, and then what those performances might do. Again, this is just an example.

If we have a fairly stable algorithm where we have fairly tight confidence intervals on training, then if we got similar training cases in and we extended those, we may see the performance bounce

around a little bit, but we would have a fairly consrained performance. We call this algorithm a more stable algorithm.

If on the other hand those error bars are very large in the training case, then it is possible with the additional training that this performance could decrease, and that is the problem that we would be interested in, as far as the FDA goes, is that this performance could be very different from what we are seeing at time zero as we get different iterations of the software in. So for both of these reasons, we are interested in algorithm stability.

In order to evaluate stability of the algorithm, we have to talk about what datasets were used in the analysis, and these are the developmental and validation datasets used by the sponsor.

There was a learning group of data, which is the training data. This was used for dimensional reduction, algorithm training, and threshold determination. There was what the sponsor has called a verification group. These were different cases from the learning group, and it was used for preliminary

verification of performance -- of algorithm performance.

Just to go through the numbers, the learning group had 65 cancer cases, 754 non-cancers, and included patients of any age. The verification group had 18 cancers, 691 non-cancers, and were limited to patients younger than 45 years of age.

Another group that the sponsors utilized was the validation group. It was an independent set of cases, and it was used by the sponsors to validate the performance of algorithm before conducting the clinical study. It was consisted of 12 cancer cases, 263 non-cancers, and again it was limited to women 45 years of age or younger.

The final type of data that was used by the sponsor was the pivotal trial data, and this was not part of the stability analysis. So I am not going to talk about this dataset here. This analysis was done on only the verification and validation in the learning group.

How can stability analysis be conducted? Well, one approach is to use bootstrap stability

analysis, and what is bootstrap sampling. This is sampling with replacement from a sample dataset, and it is a simple but powerful Monte Carlo method to assess statistical accuracy.

Just to give those who may not know an idea of what this is, if we start out with a dataset, a hypothetical dataset, of 100 cases or 100 patients, what you could do is randomly pick 100 patients from that dataset with replacement. So some of those patients would appear twice, some maybe three times, some only once, and some none at all.

You could do this under a number of iterations, and once you have those 100 different sets of data, you could do some sort of statistical analysis based on that data. So this is a statistical method that people have used to do that.

What the sponsors conducted was they estimated training variability using this bootstrap method, and they did this by bootstrapping the learning group with 100 partitions. They identified for each of these partitions the dimensional reduction. They estimated the classifier weights,

estimated the cutoff threshold for the algorithm, and then they estimated the test performance on the verification and the validation groups.

This would then allow them to -- This would then provide some estimate of the training variability associated with this device. The sponsor also estimated test variability for these same datasets. In this case, they fixed the algorithm realization using all the training data. So they selected 26 blended features. they fixed the blended feature weights, and they fixed the cutoff threshold.

Now there is just one trained algorithm. What they did was bootstrap the verification and the validation groups using 1000 partitions to estimate the test performance associated with the verification group and the validation group, and this provides an estimate of the test variability associated with the device.

So in this table I just showed the stable of T-Scan training and test variability, and I have broken it up into different columns. The first rows are shown for verification dataset, and the second for

the validation dataset.

The first row of these do different -- columns for the different rows of the training, the stability estimate associated with training or in the test, and likewise the next columns are the test. The third column is for specificity, the fourth column for sensitivity.

You can see that for the verification data that the mean estimates are roughly the same. They are in the same ballpark for specificity, with standard deviations on roughly the same order. So the variability associated with that data is roughly about the same, with about the same mean performance.

For the verification data, if we look at sensitivity for that verification data, you can see that the means now are somewhat different from each other, but again the variability associated with that verification dataset is roughly on the same order of 11 and 12 percent in standard deviation with similar confidence intervals.

For the validation data, this was the data right before preclinical when the algorithm was fixed.

We see that the specificity, again for the mean estimates of specificity, were fairly similar to each other. You can see that there is roughly a factor of 2 in standard deviation, but both of these are fairly small and modest size in specificity.

For the sensitivity, you can see basically similar matching between training and test variability. So both of these, this training variability, refers to bootstrapping the training data. Tests refers to bootstrapping the test data.

Just to try to give you a better feel for what these numbers are, what we see is I have plotted here the performance estimates with their error bars, 95 percent confidence intervals. For the bottom in purple is sensitivity, and the top, yellow, is specificity, and for the first column is for the training variability, and the second is for the test variability for the validation dataset.

You can see that the training variability is roughly on the same order as the test variability. You see a difference in the confidence intervals, but for specificity they are fairly modest for both of

these.

You see fairly large error bars associated with the sensitivity. This is most likely due, again, to the fairly small number of actual cases associated with true cancers, so that the sensitivity estimates are much -- have much larger error bars. Again, you can see that the error bars are roughly on the same order between training and test variability.

So just some comments on the stability analysis. The bootstrap analysis indicates that algorithm architecture is not unstable. However, it is also important to remember that training variability is not a trivial effect. As you could see the sensitivity error bars, those error bars are fairly big. Likewise, the error bars are a little bit bigger in the specificity estimates.

Just a reminder that the remaining speakers are going to quote only test confidence intervals based on the pivotal study data. The Panel should keep in mind that the total variability associated with the algorithm would be somewhat greater than that presented based on the test

confidence intervals.

So with that, I will end up and I will pass it off to Ron Yustein. He will give you the clinical review.

DR. YUSTEIN: Good morning. My name is Ron Yustein. I am the Deputy Director for the Office of Device Evaluation, and I will be presenting some of the clinical data this morning. And hopefully, I didn't just ruin the computer.

You have already heard the indications for use. So I am not going to go over that, and the sponsor has already gone over what the device is intended to do and what it is not intended to do, and you have seen these in your packets. So I am not going to spend the time to go over each one of these.

I wanted to talk briefly about the pivotal clinical protocols. As you have heard before, this was a two-component pivotal study consisting of a specificity arm and a separate sensitivity arm. The sponsor has gone over all the inclusion/exclusion criteria for the two studies.

So I am not going to repeat that, just

highlight here in the specificity arm that all cancers in these well women were presumed -- all cases were presumed to be cancer free. Therefore, any false positive -- any positive T-Scan was assumed to be a false positive, and any negative T-Scan was assumed to be a true negative, although no studies were done on follow-up with these women to confirm that status.

Then the difference in the sensitivity arm was the enrichment of patients, including pre-menopausal women aged 40-45.

The primary endpoint of the study was the relative probability, and you have seen this formula before where Se is the sensitivity, Rca is the prevalence, the point prevalence of cancer, and Sp is the specificity, and the criterion was that if this relative probability was 2 or greater, the endpoint would have been met.

I wanted to make a couple of comments on FDA's role in the study design issue here. This device, as the sponsor did mention in their presentation, was considered a non-significant risk device. Therefore, no formal IDE was submitted to the

FDA. However, there were several pre-IDE meetings and teleconferences held between FDA and the sponsor in the months preceding the start of the study.

I went back through the minutes of those meetings in our administrative file, and I just wanted to make a couple of points here.

FDA did express some concern over estimating the sensitivity and specificity from different study populations. However, we did agree at the end that this was a reasonable approach.

FDA did agree that it was acceptable to enrich the sensitivity arm due to low prevalence of disease. However, we did not set a limit on that and did not say what would be an appropriate level of enrichment.

We also stated that FDA would request a breakdown of the study results for those patients aged 30-39, compared to those 40-45 separately. If the sensitivity in the older patient group was less than that in the younger patient, we would probably have no major issues. However, if the opposite were true, it could present a challenge for modeling the results.

We also did agree that a demonstration of relative probability of greater than or equal to 2 would be a clinically meaningful and reasonable approach.

I also wanted to mention one item. In the sponsor's presentation they said that the criteria of 2 was set forth by FDA, in their statistician's presentation. I just wanted to emphasize that the study and the criteria were proposed by the sponsor. We agreed to them. We did not set those criteria.

Moving on to the specificity arm results, you have seen these results before. So I am not going to duplicate a lot of these. I just wanted to point out the fact that there were some variations in specificity results based on bra size, race and country of origin which did meet statistical significance.

Dr. Vishnuvajjala in her statistical presentation will go into a little bit more detail on the baseline characteristics between the two nations, the subjects in the two nations, and also the results.

Turning to the sensitivity arm results, I

wanted to -- These first two slides, I wanted to point out, because we will be asking you later in the day to comment on the degree of enrichment that took place in this study.

To highlight here, although the device is intended for patients 30-39, clearly almost 60 percent of the patients enrolled in the study as far as those contributing to the sensitivity calculation were in the age of 40-45. In addition, over 80 percent had an abnormal CBE, and about one in every seven had a positive family history.

As you will see throughout our presentation, we do break down the results between the United States and Israel. That is one of the main reasons we have come to you today, is to ask for your interpretation of the data differences between the United States and Israel.

These are the baseline demographics for those patients in the sensitivity arm who contributed with cancer, the 87 patients broken down by U.S. and Israel for four of the covariates.

I will point out here that there was a

difference, 84 percent in Israel versus 73 percent with a positive CBE, although none of these numbers up here met statistical significance.

The sponsor has already showed you the sensitivity results overall and by covariates. Just point out again that there is some variation, although the numbers are small, in bra size and based on hormone use as well, but again none of these met statistical significance.

This is a slide that you have not seen. This is some data that we looked at ourselves, looking at the line data. This may address some of the issues discussed earlier by the Panel regarding race and ethnicity.

What I have done here is broken down the sensitivity results based on country and race. The complicating factor here is that Israeli sites in the study did not necessarily record the ethnicity or race of the subject, and those accounted for the majority of patients that were in the sensitivity arm.

So, therefore, if we assume that all the Israeli patients are Caucasian, which the sponsor has

proposed to us, then these are the numbers that you get, assuming that. I kind of arbitrarily broke it down between Caucasians and then summed up the non-Caucasians, because all these numbers are very small. However, this again did not meet statistical significance.

This chart basically breaks down the sensitivity results by age and country. Up in the upper righthand corner is the 26.4 percent that you have seen the sponsor show you. if you go down this column, it breaks it down into U.S. and Israel. So in the U.S. the sensitivity was 11.5 percent, and Israel 32.8 percent.

If you move over on this side, it breaks it down by -- these columns break it down by the age, sensitivity of about 19 percent in the intended population and 32 percent in the 40-45-year-old group.

One of the things we will be asking you to discuss later today is how these results can be applied to the intended population. Just to remind you, the intended population is 30-39-year-old women who are also negative on clinical breast exam and

negative family history.

Of the 87 cancers that contributed to the sensitivity calculation, only four of those patients met all of those criteria, and the sensitivity in that group was 25 percent. Dr. Vishnuvajjala will talk about this a little bit more in her presentation.

This isn't in your packet. I just added this slide. I have this as a back-up slide, but I just wanted to point out, the statistician for the sponsor had a slide in which he showed that there was no statistical significant differences among sites with sensitivity.

I broke it out slightly different, in that I am showing that the lefthand set of slides is all U.S., and the right is Israel. You can see that most of the U.S. sites had a sensitivity of zero.

There were 131 cancer patients that were actually enrolled into the study, although the sensitivity calculations was only based on 87. So in other words, there were 44 patients who were excluded for one reason or another. This chart basically shows you the reasons for those exclusions.

Across the top, the number of patients who were excluded for that particular reason and the sensitivity of the device in that particular excluded subgroup. If all these patients are added back in, the overall sensitivity was about 23 percent.

The age limits were defined in the protocol eligibility criteria, as was prior chemotherapy was an exclusion criteria. I am just going to talk very briefly about this post-menopausal group and the technical group.

Technical issues at U.S. sites: I'm a little confused. The sponsor said that there were -- only one site had patients excluded based on technical issues. Going through the line data, I believe it is actually two sites. RFW is a site that enrolled 37 subjects. However, none of them were included in the sensitivity analysis at all. All of them were excluded for some reason or another, including 31 that were excluded due to technical reasons.

Then RJG was another site that included 56 subjects, although about 39 of them were excluded, and 34 of those 39 were excluded due to technical reasons.

In all, we lost about 19 subjects who had positive biopsy cancer from these two sites because of technical issues. The study results on the line data listed 18 -- 19 as being T-Scan negative. Then I just wanted to point out that, of those 19, six that we lost were actually in the intended patient age group.

Sensitivity results with post-menopausal women: The analyses that the sponsor presented this morning and my analyses that I have shown up until now do not include seven post-menopausal women aged 39-45 with cancer who were not included in these analyses.

Certainly, one may argue that there is a reason for excluding these subjects, namely that post-menopausal women are at higher risk and, therefore, fall outside the intended use population for this device. Also you have heard the sponsor say that post-menopausal women have different breast tissue characteristics and, therefore, it is a legitimate break point there.

Reasons for including these subjects in such an analysis: (1) the original protocol as written by the sponsor did not specifically exclude

post-menopausal women from enrollment and, in fact, 24 women were actually enrolled into the sensitivity arm that were post-menopausal, including seven with biopsy proven cancer.

Number 2, the indications for use sought by the sponsor today do not specifically exclude post-menopausal women and, third, post-menopausal women, although few, were not excluded from the specificity arm.

Dr. Vishnuvajjala will discuss more analyses based on this. I just wanted -- This is kind of a complicated slide, but I am just going to point out a couple of things.

If you do include the post-menopausal women, the sensitivity doesn't change that much here. It is 19 in age 30-39, which is the same if you exclude them, and it is 30 instead of 32 for 40-45-year-olds. Then the overall sensitivity barely changes, goes from 26.4 to 25.5 percent.

So let me just move now on to the primary endpoint. This is just a repeat of my earlier slide, just to remind you that the primary endpoint was this

calculation of relative probability, with the success criterion determined to be 2.

The sponsor has already shown you that, based on using the specificity of 94.7 for all evaluable pre-menopausal women from the specificity arm and the 26.4 from the sensitivity arm, and using an assumed prevalence of 0.15 percent, they have a relative probability of 4.95, and their primary endpoint was met.

The sponsor also provided in the submission the results for those women only age 30-39. So again, the specificity of 94.7 from the specificity arm and then using the sensitivity of women 30-39 from the sensitivity arm, which was 18.9 percent, holding the prevalence study at 0.15 percent, the relative probability is now 3.60. Still meets the 2, although here our lower bound of our 95 percent confidence interval now goes below 2.

If we add the post-menopausal women back in, things don't change very much. So now instead of the 26.4 percent sensitivity, if we substitute the 25.5 percent, we still have the relative probability

of 4.78 with the confidence interval still above 2.

Breaking down U.S. versus Israel -- sorry, this is a little bit of a complex slide, but basically this column is the United States, this column is Israel. This excludes the seven post-menopausal women. This row includes them.

What I did here was I took the sensitivity from the U.S. patients from the sensitivity arm, the specificity from the U.S. patients from the specificity arm, assuming a prevalence of 0.15 percent, and I did that for all four blocks.

What you can see here is that in the Israeli patients, regardless of whether you include post-menopausal women, the relative probability is still over 4.4, and the lower bound of 95 percent interval remains above 2. However, in the United States the sponsor did state in their concluding slides that the relative probability does still remain above 2. That is true. However, the lower bound of the 95 percent confidence interval now goes below 1.

I added this slide in on the break. I'm not sure if this answers part of Dr. Romero's

questions. I'm sorry, I don't have 95 percent confidence intervals. I pulled this out of my backup slides.

If you pull the sensitivity for the various ethnicities and the specificities from the specificity arms, these are the relative probabilities you get: 6.40 for Caucasians, 2.41 for African Americans. Again, I don't have the 95 percent confidence interval, but I would guess that it is going to be below 2 for that. Because no cancers were detected in Hispanics or Asians, the sensitivity was zero. So, therefore, we really can't calculate a relative probability, and there were no cancers within the American Indian group.

I just wanted to give my math slides here. These are -- Because the formula that was used to generate the relative probability may be a little complex, I just kind of put together three slides here, just to show you how changes in one parameter while the other two are held constant can affect your overall relative probability.

The numbers I have selected outside of

yellow are just randomly selected numbers. They have no meaning here. I didn't choose ones based on results from the study. The one in yellow is the result from the study. So here what I am just showing is that, if the sensitivity increases, you get an increase in this order of relative probability. If it decreases by 5 percent each time, you get this decrease in relative probability, holding the specificity and prevalence the same.

Likewise, if the sensitivity is held at 26.4 and we alter the specificity, you can see that the overall relative probability changes with a greater degree with smaller changes in specificity. If the specificity was 99 percent, the relative probability would be over 25, but then for every point that you drop below that, you start seeing significant changes.

Then finally, I think the sponsor's statistician also gave a slide like this. If you hold the sensitivity and specificity the same and alter the prevalence of the disease -- and again, these are just randomly selected numbers -- you can see that the

overall relative probability does not change very much.

Positive predictive value: I'm not going to go through this slide. It is in your Panel pack as to how the sponsor calculated the positive predictive value, and I think Dr. Ginor showed you at the end how they came up with a 0.734 percent, which translates into one in 136 T-Scan results being patients at risk for cancer, and that would be significant compared to what they believe is one out 167 based on the general population.

My two points here is that this one out of 136 is based on an assumption that mammography would detect 100 percent, and the sponsor did take that into account in their later slides and did show you that, if it something like 70 percent, it will go up to one in 194.

The other point I wanted to make here is that, as opposed to the previous slide where I showed you that the change in relative probability isn't affected as much by changes in prevalence, the positive predictive value is.

Just to illustrate that point, if the table that the sponsor created to adjust their prevalence and their data to a prevalence of 0.15 percent is instead -- and I'm not saying the 0.05 percent is the correct prevalence; that is one of the questions we are asking you today. This is just an illustration of how a change in prevalence can make a big change in the positive predictive value.

So if we went back and changed the prevalence to 0.05 percent, put it back into the table and redid our calculations, the positive predictive value would actually be one in 400 T-Scan cases, and again that is assuming a sensitivity of 100 percent for mammography.

Finally, I am going to just end with a couple of more slides from other considerations. Some of these have already been touched on in your discussions earlier. FDA is going to ask you later this afternoon what you believe, if any, are the impacts of false positive T-Scan results.

Number one: Is there any effect from radiation exposure from the number of women that will

be going to mammography for a false positive scan? Dr. Vishnuvajjala will be giving you some more numbers about that in a minute.

Then on the righthand side is something that I believe Dr. Romero was -- or maybe it was Ms. Mayer was discussing earlier about the impact of a positive T-Scan and a negative workup as far as on the patient's anxiety or what they may pursue later, may ask themselves was the mammogram misread or should I have another mammogram, MRI, ultrasound, whatever.

The other issue that, like I said, we have been concentrating on in our presentations, and you will see broken out in our statistical presentations, relates to combining the U.S. and Israeli data. The sponsor has given you a slide earlier today as to their reasons for pooling the data, and that will be one of our major questions that we will be asking you for assistance on determining whether that data is poolable.

Then let me just end with two slides on the additional data. There were a couple of articles in your Panel packs sent to you a few weeks ago.

These are Dr. Stojadinovic's clinical publications. I just want to state that we did not have access to the line data, and only the publications were available to us.

There was one study that you had seen from the *Journal of Clinical Oncology* in 2005 with over 1100 subjects, both U.S. and Israeli. This study did have quite a few patients, 580, under the age of 40, and here are the results broken down based on less than 40 and then the overall. The overall did include patients, I believe, into their fifties as well, although not many of those.

The points that I thought were interesting here is that the specificity here is 89 percent versus 95 percent in the pivotal arm study, and this again was a one-arm study as far as I know. Again, providing the same prevalence, the relative probability of 4.52 is obtained. However, the sensitivity of 50 percent is only based on six cancers.

Then there are two other references that I believe were provided to you, the *Breast Cancer Research and Treatment* publication. I am not going to

really comment on this, because when looking through the methods section, I noticed that many of the sites listed -- in fact, all the sites listed were the same sites that participated in the pivotal study for this PMA. So I'm not sure what subset, if any, of these results are a subset of the PMA data that is before you today.

Then the *U.S. Military Study Annual Report* that was presented to you, again currently ongoing. This was an annual report that was provided to us. Sensitivity of 33 percent, although again based on very low numbers of only three; specificity of 93 percent, and the calculated probability of 6.0 based on the prevalence of 0.15 percent.

Again, reading this article, positive for sensitivity arm was cancers or high risk lesions. I'm not exactly sure what high risk lesions -- how that was defined.

That ends my presentation, and with that I will hand it over to our statistician, Dr. Lakshmi Vishnuvajjala.

DR. VISHNUVAJJALA: Hi. I am Lakshmi

Vishnuvajjala. I am the Branch Chief of the Diagnostic Devices in the Division of Biostatistics. Dr. Harry Bushar, who actually reviewed this PMA, couldn't be here today.

I am just going to outline the pivotal clinical study. A lot of this has already been done by Dr. Ron Yustein, but I am just going to go over some of them briefly to set up the statistical analysis. You know, you have seen all of these already.

The pivotal clinical study has two different arms. The specificity arm included patients who are assumed to be normal, 30-39, CBE negative, no follow-up, and 15 U.S. sites and two Israeli sites.

The sensitivity arm has the patients who are going to biopsy, and the patients are between 30 and 45, and the biopsy result is available for all of them. We have 12 U.S. and six Israeli sites.

These are the results in the two arms, the specificity arm and the sensitivity arm. I know there was some concern expressed by computing the sensitivity from the -- computing the specificity from

the sensitivity arm, but this was in the PMA, and we are just mentioning it here. I am not really concerned about the specificity that is computed from the specificity arm.

Actually, there is a little bit more of a concern for the sensitivity that is computed from the sensitivity arm, because you actually have a spectrum bias in this population, meaning that you are computing the sensitivity or calculating the sensitivity from a group of patients that are not like the patients that actually would use the device. So I would like to say it again. We have no concerns about the specificity from the specificity arm.

You have seen this again. These are the numbers that are provided by the sponsor. We have a sensitivity of 26.4 percent. This is not including the post-menopausal women, and we have the specificity of 94.7 percent from the specificity arm, and the assumed prevalence of 0.15 percent.

What exactly does this mean? If you have 10,000 patients, we expect about 15 of them to have breast cancer. Out of these 15, four will be T-Scan

positive, and the other 11 will be T-Scan negative.

Of the remaining 9,985 women who were not expected to have cancer, about 530 will have a positive T-Scan.

Suppose we change the prevalence from 0.15 percent to 0.05 percent. Then if we look at the same 10,000 women, here you expect only five to have cancer, and out of those about one will be T-Scan positive, and the other four will be T-Scan negative.

Out of the remaining 9,995 women, you have about 531 to be T-Scan positive. They do not have cancer.

Actually, let me go back to that. One of the things in this -- I haven't come to it, but you will see in other presentations, when you go from a prevalence of 0.15 percent to 0.05 percent, the ratio of false positives to true positives changes by a factor of four, but if you remember the calculations for what is called the relative probability, which is the primary endpoint, it hardly changes. It goes from 4.95 to 4.97.

This is just a different -- one more way

of looking at the results about what happens when the prevalence changes. Even though the primary endpoint stays constant, your false and true positive rates do change.

These are the baseline characteristics in the specificity arm. We have -- Again, you have seen all of these before. I just want to mention, some of the baseline characteristics are significantly different in the two populations and in U.S. and Israel.

These are the results in the -- the T-Scan results and the specificity in the various groups in the specificity arm.

Again, we see that the proportion -- the specificity between the two groups, Israel and the U.S., does change in almost every -- is different in almost every category that we look at, the family history, the bra cup size, the hormone use, the post- and pre-menopausal, and the overall.

These are the baseline characteristics for the sensitivity arm, which includes the post-menopausal women, and these are just the benign cases.

If you notice, the numbers are much, much smaller than in the specificity arm, but we see some of the factors are still significantly different between the two countries.

In particular, the clinical breast exam negative proportion is quite different. It is only 31 percent in the U.S., and 54 percent in Israel, and the proportion of pre-menopausal women is different also, even though it is not quite as dramatic as the clinical breast exam.

These are the baseline characteristics for the sensitivity arm in the malignant cases. We have 29 in the U.S. and 65 in Israel, and again this includes post-menopausal women. I think, out of the 29 in the U.S., three are post-menopausal, and out of the 65 in Israel, four are post-menopausal.

In comparing these two, I also like to mention, when we look at the U.S. and Israel, it is not really subgroup analysis. We routinely look at site differences when we have a multi-center study, and when we have foreign data, we also look at the differences between the U.S. population and the

foreign data to make sure that the foreign data does actually support the device in the U.S. population.

These are the proportions of cancers for women age 30-39 broken down by the CBE and family history status. Even though the slide says it includes post-menopausal women, this is because FDA has included post-menopausal women in all the calculations, but in this particular slide they are all pre-menopausal women. There are no post-menopausal women in the 30-39 that have cancer. So all the cancers in the post-menopausal women happened in the 40-45 age group and not in the 30-39.

Again, this is again the T-Scan sensitivity for women age 30-39 in the sensitivity arm. The first column, which is highlighted, is the intended use population. These are the women who are 30-39 who are negative on both the clinical breast exam and the family history.

If you go across the top row for U.S. sensitivity, it shows the sensitivity in the U.S. population in these groups. As you can see, in the 30-39 age group not a single cancer was T-Scan

positive in the U.S.

This is the sponsor's primary endpoint, which is called the relative probability. Note that this relative probability is actually the positive predictive value multiplied by the prevalence. So the relative probability is actually the positive probability -- the positive predictive value divided by the prevalence.

The success criterion proposed by the sponsor is that the relative probability be greater than or equal to 2. As opposed to the relative probability, the odds ratio is the ratio of odds of having a malignancy in the T-Scan positive group to that of the T-Scan negative group.

This can be shown to be mathematically greater than the relative probability, but this is more amenable to the statistical calculation. So we used the odds ratio in the calculations, but whatever values we got for the odds ratios, the primary endpoint, relative probability, will always be less than what we got for the odds ratio.

So in order to estimate the effect of all

these correlates on the incidence of cancer, we included these five factors into the logistic regression, and this is the order in which these factors entered the regression. Each one that comes -- Each one of these will explain the variation after the first one is accounted for.

So the family history has an odds ratio of 3.6. After accounting for the effect of the family history, the country has an odds ratio of 4.5. After accounting for those two, then you go to the third one and so on. We also looked at the confidence intervals.

We always look at a hypothesis test or confidence intervals, and we do not just accept point estimates as indicator self-effectiveness. So this is pretty routine also. If we do not do a hypothesis test for the estimates, we always have confidence intervals for the estimates.

So what this regression shows is either being T-Scan positive or being post-menopausal or not using hormones has an odds ratio with a lower bound of less than 2. Only being in Israel or having a family

history has a lower bound on the odds ratio of better than 2.

So controlling for each of these potentially significant covariates is necessary to properly assess the residual effect of t-Scan to predict malignancy.

We also did the logistic regression for the women in the 30-39 age group. This is over all women 30-39 in both Israel and the U.S. And again we see that the family history has an odds ratio with a lower bound greater than 2.6, and being in Israel also has a lower bound of being -- a lower bound of greater than 2, and for the other two, T-Scan and the hormone use, the lower bounds are pretty close to 1.

The sponsor's primary endpoint does not allow for all the effects that are competing in the prediction of the cancer, and combining the -- One of the problems in combining the U.S. and the Israeli data is for the primary endpoint -- I do not have the site in this presentation -- the relative probability is actually greater for either U.S. or the Israel -- Actually, the other way around: The relative

probability is higher for the combined population than it is for either Israel or the U.S. This actually has been in some of Dr. Yustein's slides, but they were on different slides, and I don't know if all of you have noticed that.

So we have something a little bit of a paradox going on here. You would expect the measure to be somewhere between the two countries. Instead, it is greater than either one. I think one of the reasons is it is very little affected by the prevalence unless it becomes very, very large.

The separation of the sensitivity and specificity into two arms under different protocols -- it may have unintentionally complicated in how they can be consolidated into one analysis.

So the RP may be -- It is on an entirely different scale. It may not be related to the intended population of the clinical breast exam and family health -- not family health, the family history, negative women in the age group of 30-39. Thank you.

The next speaker is Dr. Roselie Bright,

who is going to discuss the benefit risk analysis.

DR. BRIGHT: I am going to talk about the benefit risk analysis of T-Scan.

FDA and the sponsor agreed that the primary endpoint would be a relative probability of greater than 2. The sponsor met this endpoint in the unadjusted analysis.

FDA believes that it is also important for the Panel to consider the benefits and risks of this device from different perspectives. This presentation will present alternative analyses for considering the risk/benefit of this device.

First I will talk about the sponsor's benefit analysis, the underlying assumptions, FDA's calculations, and a discussion. Then I will talk about the FDA's benefit/risk analysis, the method, underlying assumptions, calculations, and discussion.

There were three assumptions underlying the sponsor's benefit analysis. The first was the prevalence of cancer. FDA verified the sponsor's conclusion that prevalence estimates between 0.00017 and 0.0015 do not affect the calculations.

The second assumption was the estimate of T-Scan specificity. FDA used the overall estimate of 94.7 percent and the low estimate of 88 percent which was found for African Americans and Hispanics in the specificity arm of the study.

The third assumption was the estimate of T-Scan sensitivity.

FDA repeated the sponsor's benefit analysis by using four different values for T-Scan sensitivity. The first value was 26.4 percent which was used by the sponsor and for all women aged 30-45, both clinical breast exam positive and negative, and from both the U.S. and Israel.

The second value was 10.3 percent which FDA calculated for all U.S. data, including 40-45-year-olds and clinical breast exam positive patients.

The third value was 5.6 percent which was for all women age 30-45 who were family history negative, and only from the U.S.

The final value was zero percent, which was for women 30-39 who were clinical breast exam negative, and only from the U.S.