

U.S. FOOD AND DRUG ADMINISTRATION  
NATIONAL INSTITUTES OF HEALTH  
CENTER FOR DEVICES AND RADIOLOGICAL HEALTH

Meeting of:

MEDICAL MICROBIOLOGY DEVICES PANEL  
COMMITTEE

~~March~~ May 20, 1999

9200 Corporate Boulevard  
Rockville  
Maryland

Reported By:

CASET Associates  
10201 Lee Highway, Suite 160  
Fairfax, Virginia 22030  
(703) 352-0091

## TABLE OF CONTENTS

	<u>Page</u>
Opening Remarks, Introduction of Panel - Dr. Charache	1
<b>Premarket Notification Submission:</b> Digene Corporation Hybrid Capture CMV Nucleic Acid Hybridization Assay	
Manufacturers Presentation - Mark Del Vecchio - Larry Kricka, <u>PhDPh.D.</u> - James Lazar, <u>PhDPh.D.</u> <del>Thomas Quinn, M.D.</del> Allison Cullen - Jonathan Kahn, <u>EsqEsq.</u>	5
FDA Presentation - Prasad Rao, <u>PhDPh.D.</u>	39
Open Committee Discussion and Recommendations	50
<b>Premarket Approval Application:</b> Gen-Probe Amplified Mycobacterium Tuberculosis Direct (MTD) Test	
Sponsor Presentation - Glen Frieberg - Vivian Jonas - Kathie Smith, PhD - Antonino Catanzaro, M.D. <del>Gail Woods, M.D.</del>	78
FDA Presentation - Patricia M. Simone, M.D. - Roxanne Shively - John Dawson	129
Industry Response	154
FDA Response	161
Open Committee Discussion, Recommendation and Vote	162

**PANEL MEMBERS:**

**PATRICIA CHARACHE, MD, Chair,** Johns Hopkins University  
School of Medicine

**MARGARET R. HAMMERSCHLAG, MD,** State University of New  
York Health Science Center

**NATALIE L. SANDERS, MD, MPH,** Southern California  
Permanente Medical Group

**CARMELITA U. TUAZON, MD,** George Washington University  
Hospital

**MELVIN P. WEINSTEIN, MD,** Robert Wood Johnson Medical  
School

**MICHAEL L. WILSON, MD,** Denver Health Medical Center

**DAVID W. GATES, PhD, Industry Representative.** Becton  
Dickinson Microbiology System

**STANLEY M. REYNOLDS, Consumer Representative.** Bureau of  
Laboratories, Pennsylvania Department of Health

**Microbiology Devices Panel Consultants**

**PAUL H. EDELSTEIN, MD, Temporary Voting Member.**  
University of Pennsylvania Medical Center

**RICHARD O'BRIEN, MD,** CDC, Division of Tuberculosis  
Elimination

**L. BARTH RELLER, MD,** Duke University Medical Center

**STEVEN C. SPECTER, PhD,** University of South Florida  
College of Medicine

**Guest Speaker:**

**PATRICIA M. SIMONE, MD,** CDC, Division of Tuberculosis  
Elimination

P R O C E E D I N G S (10:01 a.m.)

**AGENDA ITEM: Opening Remarks, Introduction of Panel.**

Dr. Charache ~~MS. POOLE~~: We will begin the meeting of the Microbiology Devices Panel. I am pleased to see old friends here as well as old friends.

We are going to start by introducing the panel. I think we can begin. Steve, do you want to introduce yourself?

DR. GUTMAN: I am Steve Gutman. I am the director of the division of clinical laboratory devices, which is the division sponsoring this panel meeting.

DR. EDELSTEIN: Paul Edelstein, University of Pennsylvania. I guess I am a voting member of the panel today.

DR. TUAZON: I am Carmelita Tuazon from the George Washington University Medical Center.

DR. WILSON: Mike Wilson, from Denver Health Medical Center.

DR. CHARACHE: I am Patricia Charache from Johns Hopkins.

MS. POOLE: Freddie Poole, I am the Executive Secretary as-exee\_.

DR. HAMMERSCHLAG: Margaret Hammerschlag from

SUNY Health Science Center in Brooklyn.

DR. SANDERS: ~~Madeline~~Natalie Sanders, Southern California Permanente Medical Group, also known as Kaiser. We are in the Harbor City Los Angeles Medical Center.

DR. WEINSTEIN: I am Mel Weinstein from Robert Wood Johnson Medical School, New Brunswick, New Jersey.

MR. REYNOLDS: Stan Reynolds, Pennsylvania Department of Health, Bureau of Laboratories.

DR. GATES: David Gates of Becton Dickinson. I am the industrial representative.

DR. RELLER: Barth Reller, Duke University Medical Center.

DR. SPECTER: Steven Specter, University of South Florida College of Medicine.

DR. CHARACHE: Thank you very much. I think Freddie Poole will now disclose any conflict.

MS. POOLE: Good morning. For the record, the following statement addresses conflict of interest issues associated with this meeting, and is made part of the record to preclude even the appearance of an impropriety.

To determine if any conflict existed, the agency reviewed the submitted agenda and all financial interests reported by the committee participants.

The conflict-of-interest statutes prohibit special government employees from participating in matters that could affect their, or their employees, financial interests.

However, the agency has determined that participation of certain members and consultants, the need for whose services outweighs the potential conflict of interest involved, is to the best interests of the government.

We would like to note for the record that the agency took into consideration certain matters regarding Dr. s. Paul Edelstein, Margaret Hammerschlag, Barth Reller, Melvin Weinstein and Michael Wilson.

These panelists reported current and/or past interest in firms at interest on matters not relating to what is being discussed today.

Since these matters are not related to the specific issues of this meeting, the agency has determined that these panelists may participate in today's deliberations.

We would also like to note for the record that the agency took into consideration a matter regarding Dr. Richard O'Brien, who reported his institution's involvement in a related matter with a firm at issue.

The agency has determined that, because this interest is imputed to him, Dr. O'Brien may participate in the mycobacterium tuberculosis discussion.

In the event that the discussions involve any other products or firms not already on the agenda, in which the participant has a financial interest, the participant should excuse him or herself from such involvement, and the exclusion will be noted for the record.

With respect to all other participants, we ask in the interest of fairness, that all persons making statements or presentations disclose any current or previous financial involvement with any firm whose products they may wish to comment upon.

For today's meeting, Dr. Paul Edelstein was appointed as temporary voting member.

For the record, he is a special government employee and is a consultant to this panel under the Medical Devices Advisory Committee.

He has undergone the customary conflict of interest review, he has reviewed the material to be considered at this meeting. It is signed, Elizabeth E. Jacobson, PhDPh.D., acting director, Center for Devices and RadiologieRadiological Health, May 10, 1999.

Before we get started, for old business, the panel last convened on February 11, 12, and 13. At all three meetings, the panel recommended that guidance documents be developed to help to address the issues, and those guidance documents are in the process of development. Thank you.

DR. CHARACHE: Thank you very much. Does anyone have any questions for Freddie Poole?

I think we can begin the business of the meeting.

The premarket notification ~~and~~-submission is our first discussion, Digene Corporation Hybrid Capture CMV Nucleic Acid Hybridization Assay for the Chemiluminescent Detection of Cytomegalovirus (CMV) DNA in White Blood Cells.

The focus is the question of terminology, whether it is appropriate to call this technique the use of signal amplification technology.

We will begin with the manufacturer's presentation, Mark A. Del Vecchio.

**AGENDA ITEM: Premarket Notification and Submission. Manufacturer's Presentation.**

MR. DE VECCHIO: Good morning. As Dr. Charache indicated, I am Mark Del Vecchio. I am the associate

director of regulatory and clinical affairs at Digene Corporation.

I would like to thank the members of the Microbiology Advisory Panel meeting and the Division of Clinical Laboratory Devices for giving Digene the opportunity to present this morning.

I will be leading off a series of discussions directed toward describing the review history of the hybrid capture CMV assay, its relevance to the discussions regarding signal amplification, and the appropriateness of that term.

As indicated, Digene has requested that the CMV DNA assay reflect signal amplification to describe the technology that is utilized.

Broadly defined, the CMV DNA assay detects CMV DNA in blood samples from immunocompromised patients.

These are some of the characteristics of the assay, including solution hybridization, antibody capture, signal amplification, which is the subject of this morning's presentation.

Before discussing the details specific to the CMV review, it is important to note that Digene and DCLD have used the term "signal amplification" to describe the technology on several occasions related to other Digene

products, specifically, the chlamydia trachomatous assay, the neisseria gonorrhoea test, as well as the hybrid capture HPV DNA assay.

In fact, in the summer of 1997, the signal amplification terminology was submitted and approved in a PMA supplement for the hybrid capture HPV test.

For more specifics regarding the CMV DNA assay and its review history, just prior to obtaining clearance for the CMV test, in the final stages of the review, Digene and DCLD reviewers engaged in lengthy discussions regarding the use of the term "signal amplification" in describing the device methodology and test principles.

Those discussions included peer-reviewed journal literature review, analytical and clinical sensitivity data, and discussions related to the appropriate degree of amplification that distinguishes the technology as signal amplification.

With respect to these earlier discussions, during which the term "signal amplification" was used, and familiarity with the use of that term, surprisingly, DCLD determined that Digene could not use signal amplification in its labeling.

After discussing several alternatives, DCLD proposed the use of the term "signal enhancement" to

describe the hybrid capture technology.

In order to obtain a timely 510(k) clearance, Digene reluctantly chose to accept the use of the term signal enhancement instead of signal amplification.

We maintain that the use of signal enhancement does not accurately or scientifically describe the Digene technology.

Moreover, the use of signal amplification applied to the technology is not only described in peer-reviewed literature, but amongst the scientific community.

Therefore, Digene believes strongly that we should be permitted to use the term "signal amplification" to describe our technology in the product labeling.

In the earlier information that was forwarded to you prior to this meeting, there were several alternatives offered for a description of the Digene test.

One of those descriptions, nucleic acid signal amplified solution hybridization, Digene believes to be the appropriate choice amongst those alternatives, and we would be prepared to include that terminology to describe the CMV assay, in our product labeling.

In order to support, to provide scientific support for the use of that terminology, I would like to introduce Dr. James Lazar, the director of the clinical sciences group of Digene's research and development department.

Dr. Lazar has been with Digene almost 10 years now, and he was integral in the development of the hybrid capture system and was the project leader for the development of the CMV DNA test.

DR. LAZAR: Thank you. What I would like to start out doing this morning is reviewing some of the printed material that is currently available that is relative to the question of signal amplification.

I will start out by reviewing what FDA has said about signal amplification, reviewing a document from the NCCLS, and reviewing some of the peer-reviewed literature that describes hybrid capture signal amplification.

I will then give the podium over to Dr. Larry Kricka to talk about the definition of detection methodologies. Then I will come back and talk a little bit more about the analytical sensitivity and clinical sensitivity, the way to do the CMV assay.

The FDA's review criteria for nucleic acid amplification-based diagnostic is not a specific

definition for signal amplification, but FDA has recognized that these types of assays do exist.

As stated, additional techniques may amplify the intensity of the detectable signal of a hybridization reaction. However, no other details are offered in this guidance.

The NCCLS has given a little more specific definition, in that the use of specific detection methodologies may directly increase the signal in proportion to the amount of target in the reaction.

They use the branched DNA as an example of signal amplification.

Just recently, in a reviewed article of essentially a tutorial of molecular amplification methods in infections in medicine, hybrid capture was specifically described as a signal amplification system, and listed in a table with branched DNA as one of two signal amplification systems that are available today.

It is not just recently that this term has been used. It has been used as early as 1994 with reference to hybrid capture, describing our hepatitis B assay that we currently market in certain export markets, as a quantitative, non-radioactive hybrid capture assay, coupled to sensitive chemiluminescent signal

amplification system.

In an evaluation involving researchers at the NIH, and our HPV-based hybrid capture assay, they specifically describe hybrid capture as a signal amplified test, and specifically differentiate it from a DNA amplification method, such as PCR, and a non-amplified method such as a southern or a dot blot.

In a recent review on CMV detection methodologies, Michael Boeckh and Guy Boivin again describe this as a solution hybridization assay, involving amplified chemiluminescent detection.

In a letter to the editor, Yi-Wei Tang and David Persing, from the Mayo Clinic, also described hybrid capture as a signal amplified probe technique, although we don't agree with the rest of this comment. That is a correct reference to hybrid capture as a signal amplified probe technique.

Again, there are other articles in the peer-reviewed literature that I have not covered and there is also literature that is not peer reviewed -- industry magazines, trade journals, that also recognize and have described hybrid capture as a signal amplification technology.

We have not been able to find, nor has FDA been

able to provide us with, an example in the literature that actually refutes the description of hybrid capture as signal amplification technology.

I would like to turn the podium over now to Dr. Larry Kricka. Dr. Larry Kricka is a professor of pathology and laboratory medicine at the University of Pennsylvania Medical School.

He is also director of the general chemistry laboratory, and is a recognized expert in immunoassay theory design and analysis.

Dr. Kricka will address the concept of defining detection method technologies.

DR. KRICKA: Thank you very much. I would like to start by making a statement that I am receiving fees and expenses in connection with this presentation from Digene. Otherwise, I have not received, currently or in the past, any financial payments from Digene, or had any other sort of relationship with them.

There are a number of different types of assays that have been developed for nucleic acids. They range from simple, direct probe assays, to ones which employ signal amplification, probe amplification or target amplification.

I am going to concentrate, focus, on signal

amplification as it applies to the hybrid capture assay.

Unfortunately, in this area, definitions have lagged behind development. I would suggest a feature analysis method which will provide a framework for distinguishing different sorts of assays.

This will, then, naturally lead to a proposed definition of signal amplification.

I think the best place to start is to look at some of the components of the detection system. First of all, there is a probe which provides the molecular recognition of the targets.

One thing to bear in mind here is that the size of the probe can vary. There is a minimum size that is required for specificity, but in some assays -- as we will see in a minute -- you can lengthen the probe or increase the number of probes that are used, specifically to gain increases in the amount of signal.

After the recognition event, there needs to be some signalling event, and this can either use a label of some sort -- that may or may not employ further reagents as part of a detection system.

Now, I am going to categorize the detection system using these three features: the number of recognition events, the number of labels there are per

probe that is used in that recognition event, and then the number of signals that you can generate from the labeling.

Each of these factors are a way in which you can increase the amount of signal that you get out of your assay tube.

The interesting thing about this is that they are not simply additive, although they can become exponential in terms of the number of signals that you finally get from the assay tube.

Let me just illustrate this with a simple, non-amplified detection assay that shows these three different attributes: the recognition event, the label and the signal.

Here is the assay over here, where a single DNA probe is detecting a single DNA molecule. It is attached to a single label and the label is giving a single signal.

So, this is the simplest type of system in terms of these three features of the assay.

Now, obviously, you can take each of these features and begin to improve them, with a view to improving the amount of signal that you get out of your assay tube.

Here is another example. This is an ELISA. At the bottom we see the ELISA design, where an antibody has captured a single antigen, and a single monoclonal antibody is bound to it.

Attached to that is a single label, but this time this is an enzyme molecule.

Now, this is a multiplying label, because one enzyme molecule will produce many, many signals. So, in this particular type of assay, we still have one recognition event, we have only a single label. But in this particular category here, we are getting more than one signal from that enzyme labeling.

Now, it is possible to look at these three different features of an assay and try to maximize each of them to gain the biggest advantage in terms of signal, and eventually get an exponential increase in the signal from the assay.

Here is the first example of one of those designs. This is a branched DNA assay, in which we have multiple recognition events, multiple labels and multiple signals from the individual labels.

So, in the assay design, if you start here, here is the single molecule with the analyte, and more than one probe is detecting this, as the recognition event.

Each of these probes, in turn, is attached to a label system which carries on it many labels. Because these are enzyme labels, they each produce multiple labels.

The net effect is an exponential increase in the signal, shown over here, as a result of the combination of multiple recognition events, multiple labels, and multiple signals per individual label.

Another assay design achieves the same effects, and this is the hybrid capture assay.

This is an antibody on the solid support. This is the target and this is the RNA probe which is binding to the target.

The first thing to appreciate here is that this probe is very, very long. This is a first 9,000-base-pair probe that is being used in this assay.

In terms of the amount of probe that you would get for specific detection, there is obviously much, much more than you might want.

In fact, if this was a human genome, if you wanted unique sequence on a probe that would be unique, you only need a 17.

So, the purpose to all this extra probe is, in fact, to provide amplification sites. In the assay, an

antibody which is labeled with several enzyme molecules binds to this duplex.

So, the assay involves multiple recognition events provided by these antibodies, recognizing this duplex. Each of these antibodies has more than one label. Because they are enzyme labels, they each produce more than one signal.

The net effect in this is that we have all three factor acting in concert to produce an orders-of-magnitude increase in the signal.

I just summarized these two assay designs here, to show their similarities.

Recognition events in branched DNA, there are many probes. In hybrid capture, there is the equivalent of many probes because, along that length of RNA probe that forms its duplex, you have over 1,000 sites on which the antibody is combined.

Labels, both have more than one label attached to the probes. Then, both because they are enzyme labels, you have many signals, the net effect is this orders or magnitude increase in signal.

So, this is a summary of the different types of assay. What I have done is, I have compared the two extremes in assay design based on the features that I

have chosen -- recognition events, the number of labels on the probe, the number of signals per label that is on the probe.

At one extreme, we have this simple, non-amplified assay where it is one recognition event, one label on the probe, one signal per label.

That could be a simple southern blot with a fluorescent label.

At the other end of the spectrum are the assays in which each of these factors is increased, where you have multiple recognition events by multiple probes. The probes have multiple labels on each probe, and the labels themselves are capable of giving multiple signals per label, and the net effect is this enormous increase in signal.

This leads, then, naturally to a proposed definition of the signal amplification assays. This is put up on this slide, an assay format that incorporates multiple recognition events and multiple labels, to increase the measured signal by orders of magnitude above a simple, one label, one probe, one binding event design.

I will hand over the podium now to my colleague, who will deal with this technology in detail.

DR. LAZAR: Thank you, Dr. Kricka. That was

very informative.

What we would like to do now is to show you a little more about hybrid capture in detail. Because it is very difficult to visualize what is going on in the hybrid capture reaction itself, we have put together a little three-dimensional, computer-generated video, which I will try to talk you through now.

I think it will help you get a better feeling for what is actually going on at the molecular level.

The first step in the assay is the denaturation of the target DNA. Most DNA are double stranded and denaturation is necessary to separate the double strands of DNA and make them available for hybridization.

Digene uses a concentrated base solution coupled with heat to achieve this step. The base solution also lyses the cell and liquifies the protein and prepares the sample for hybridization purposes.

Following denaturation, an RNA probe is added and, as Dr. Kricka mentioned, for CMV it is a 39,000 base pair probe.

It is a single stranded, unlabeled RNA probe. This hybridizes to the single stranded DNA, forming an RNA:DNA hybrid.

Again, for each copy of CMV that is in the

sample, 39,000 base pairs of RNA:DNA hybrid are formed and detected.

After the hybrid is formed, it is transferred to a capture, a solid phase capture phase, where the hybrid is captured with an antibody specific for RNA:DNA hybrids.

This solid base can either be a coated tube, as it is for the CMV assay, or it can be a coated microplate, as it is for some of our other assays.

After the capture step, the antibody conjugate is added. Mark is going to pause it right there, so you can really see what is happening there.

Again, you are really only looking at a very small part of the target of just one molecule. These antibody targets are coming in here, and three-dimensionally wrapping up this hybrid, covering it up completely.

As you can see now, we can get a much better picture of what is really happening in solutions that have hybrid, and how many antibodies are being bound to it.

Once it is detected with the antibody, of course, there is a washing step, and a chemiluminescent substrate is added.

The alkaline phosphatase on the conjugate antibodies cleaves a phosphate group from the substrate molecule and releases a photon of light.

The photons of light are counted in a lumenometer. The intensity of the light emitted, then, denotes either the presence or the absence of the target DNA.

I think you can get kind of a good sense of what is happening there in hybrid capture.

Now, in order to estimate what the amplification fold, or an amplification factor is for the hybrid capture, I say it is necessary to use some hybrid capture facts.

One of these is that the antibody footprint -- that is, the minimum size of RNA:DNA hybrid that can be detected -- is somewhere between 10 and 20 base pairs.

On average, the conjugate, the anti-RNA:DNA conjugate molecule, contains 2.5 alkaline phosphatase molecules. That is, some of them will contain two and some of them will contain three.

To accommodate sites for binding in the solid phase, we estimate that one conjugate combines approximately every 30 base pairs.

I will define these numbers to the hybrid

capture CMV assay which, again, has 39,000 base pairs of probes. One can calculate an amplification factor, or essentially, fold the amplification above what you would have if you just had one binding event, one probe, and one signal in molecule.

The way that we can do that is to use those numbers, dividing 39,000, which is the length of the RNA:DNA hybrid, dividing it by 30, for each conjugate molecule, and multiplying by 2.5, which is the number of alkaline phosphatase molecules per conjugate.

That gives you an amplification factor of over 3,000-fold.

Just to a little bit more further develop Dr. Kricka's idea about what the minimum dose size is necessary for specific recognition of a target being a 17-base pair sequence, there would actually be about 2,300 of those sequences within the RNA:DNA hybrid that is being detected in this CMV assay.

Again, this is very different from an ELISA-type assay. We are actually forming the antigen in the hybridization reaction.

We have control over how big that hybrid is. We have 39,000 in the CMV assay, but that is only 17 percent of the genome.

If we needed more sensitivity, we can always add more probes. We can show that the sensitivity is directly related to how much probe you have.

If you have more probe, of course, you are going to bind more target, you are going to bind more copy to molecules, and you are going to increase the amount of signal that is produced.

In the analytical sensitivity of the hybrid capture assay from a whole blood sample is approximately 700 copies per ml.

There have not been well-controlled comparative studies yet between hybrid capture and other nucleic acid detection technologies for CMV, since these other assays -- bDNA and AmpliCor PCR -- have not been cleared for CMV detection.

However, in the same review of CMV methodologies, it was reported that the bDNA assay had a sensitivity of 900 copies per million leukocytes.

So that we can get the units roughly equivalent here, immunocompromised patients typically have somewhere between one and five million leukocytes per ml of blood.

The AmpliCor PCR assay is from a plasma sample, and one ml of plasma may come out of 2.5 to 3 mls of blood, and had a reported sensitivity of 1,000 copies per

ml.

So, clearly, there is no significant difference in analytical sensitivity between the hybrid capture CMV assay, the research version of the branched DNA CMV assay, and the research version of the Amplicor PCR target amplified assay.

So, how does it perform clinically, because that is really the key issue, and in a multi-center clinical trial that was just published in the Journal of Clinical Microbiology in April of this year, there was a report of a multi-center trial of this assay in an immune compromised population.

This included HIV and AIDS patients, bone marrow transplant patients and solid organ transplant patients.

In this study, hybrid capture was compared to antigenemia and shell vial culture and traditional sputum culture, or compared to a consensus positive result.

In the HIV/AIDS population, hybrid capture was 93 percent sensitive, versus 47 percent for shell vial, and 55 percent for traditional culture.

In the bone marrow population, hybrid capture was 92 percent sensitive, while shell vial and culture were approximately 70 percent sensitive.

In the solid organ transplant population, which

really is the largest population for the use of this type of assay currently, hybrid capture was 97 percent sensitive, while shell vial and culture were both less than 25 percent sensitive.

Clearly, it is clinically superior to the traditional shell vial and culture methods.

These results were recently confirmed, just last week, in presentations at the Clearwater Clinical Virology Symposium where, among a number of abstracts, two that I have listed here -- one from the Cleveland Clinic -- again compared the hybrid capture assay to renal transplant patients to traditional culture. They achieved 100 percent sensitivity and 99 percent specificity.

In a different study, hybrid capture was compared directly with shell vial and, again, achieved 100 percent sensitivity and 100 percent specificity, while shell vial only achieved a sensitivity of 60 percent for clinically indicated CMV pneumonia.

Now, because there is no clear comparison method between hybrid capture and a target amplified CMV assay or another signal amplified CMV assay, we invited Dr. Thomas Quinn from Johns Hopkins to come and talk to you today about his experience with our chlamydia assay,

which is also pending clearance through FDA.

Unfortunately, Dr. Quinn was not able to attend today. So, in his absence, I would like to introduce Ms. Alison Cullen, again, from Digene's research and development department. Ms. Cullen will actually give Dr. Quinn's presentation.

MS. Cullen has been a member of Digene's R&D team for at least 10 years, and she has been, again, integral in the development of the hybrid capture technology, and has been a team leader in the development of the chlamydia and gonorrhea assays of Digene, and has worked collaboratively with Dr. Quinn.

MS. CULLEN: Thank you, Jim, and good morning. I would like to take a moment to discuss the relevance of chlamydia trachomatous detection to the research questions.

As Jim has mentioned, Digene has developed a test for chlamydia trachomatous detection that is based on the same hybrid capture technology used in the CMV test. This test is current pending FDA approval.

As many of you are aware, there are a number of FDA-cleared commercially available nucleic acid based tests for CT detection, and a number of them are shown in this slide.

This gives us the unique opportunity for comparing the hybrid capture system to these other tests. We did not have this opportunity for CMV.

There are some parallels between CMV and chlamydia detection, in that both use tissue culture as a gold standard and, in recent years, nucleic acid based detection has been shown to be more sensitive than culture.

All these test methodologies have been evaluated over the years in Dr. Quinn's laboratory. That included the Digene hybrid capture test.

If you were to look at the characteristics of each of these assays, you would say that culture uses a multiplication of the organism, and no amplification, in order to detect the organism.

Gen-Probe's PACE 2 assay, the target is the multiple copy, but there is no amplification involved.

The remainder of these assays all use multi-copy targets, and a variety of different amplification methods.

If you were to review all the package inserts for these assays, you would be able to look comparatively at the sensitivity. All these sensitivities are compared to culture.

In this slide also is shown a 100 percent sensitivity, which we know is not the case, but it does give us a relative comparison.

On this comparison, you can see that the hybrid capture test has excellent clinical sensitivity, and this sensitivity is comparable to some of the target amplified tests.

I wanted to talk about a study that was performed in Dr. Quinn's lab on a subset of specimens from the multi-center clinical trial.

This study was recently published in a peer-reviewed journal, the Journal of Clinical Microbiology. I am the second author on that paper, and it is available for you to review after the presentations.

It is interesting to note, in this publication, that the hybrid capture technology is referred to as a signal amplification-based test.

Now, 587 patient specimens were analyzed in this study, and they were collected from Baltimore STD clinics.

All the specimens were tested by all three methods -- culture, hybrid capture and the Amplicor PCR test.

In addition, once they would test positive,

culture negative specimens were resolved by nucleic.

This table summarizes the results. You can see from these data that the hybrid capture tests have excellent clinical sensitivity and specific as compared to Amplicor PCR.

There is 95.4 percent sensitivity for hybrid capture, and Amplicor PCR had 90.8 percent sensitivity. Both of these tests, the Amplicor tests, had significantly more sensitivity, statistically significant more sensitive than culture, which had 81.5 percent sensitivity.

From this study, you can conclude that the hybrid capture test, which is an example of signal amplification, and the Amplicor PCR test, which is an example of targeted amplification, demonstrated statistically lower clinical performance, and that both tests were significantly more sensitive than culture.

I didn't have time to present some other information on the studies that we have performed with the hybrid capture test on video, but we have done those tests through other target amplified tests such as LCR, and those are in preparation for publication.

The same comparable results to the study that I just presented were found.

So, overall, we can conclude that the Digene test is significantly more sensitive than tissue culture, traditional ELISA and direct probe tests, and that it shows equivalent clinical performance to the available target and probe amplification detection system, such as PCR or LCR.

These data further support our use of the term signal amplification in describing all of our hybrid capture tests. Thank you very much.

MR. KAHN: My name is Jonathan Kahn. I am a partner in the law firm of Hogan and Hartson, and outside regulatory counsel to Digene.

My only financial interest is, of course, having my modest fees paid by Digene every 30 days, hopefully.

Digene asked me to try to add some perspective to what is clearly a very unusual situation. I can't remember -- I have been doing this about 25 years.

Although I have helped lots of companies -- dozens of companies -- at panel meetings, I rarely speak.

They have asked me to speak today, primarily to try to put into perspective FDA and the company bringing to you a company related to a 510(k) labeling question. That is not typically what you hear.

It really is even more unusual, in that we are

talking about a question of the proper terminology to apply to the Digene technology.

Digene is strongly of the view that the technology should be described as a signal amplified technology based upon, as you have heard, the science and the literature.

They believe that it is not only applicable to CMV, it is also applicable, as you might have concluded, to all of our hybrid capture technology.

Digene and FDA have been discussing this issue in some detail, and I think there is a good faith disagreement between FDA and the company.

I believe that FDA has a legitimate concern as to where they draw the line between the degree of amplification that should be required for describing an assay as signal amplified.

What we do believe here, however, is that FDA drew the line improperly with regard to CMV and the hybrid capture technology.

The agency asked Digene to give in and agree to signal enhancement. It is the company's position that, if anything, signal enhancement is not the appropriate terminology, primarily because there is actually no basis in either the literature and science for utilizing that

term in connection with this technology.

Why does Digene really care about this? One might say this is a simple labeling issue, it is a 510(k). Why are we all here spending our time talking about this, when your time is very valuable.

The answer is that signal amplification has become a well-recognized term, in both the industry, in science, in physician use and in laboratory use.

The company strongly believes that it would prejudice Digene if its technology was not allowed to be described as signal amplified.

It would, in the company's view, prejudice those who use the technology to believe that, somehow, this is a lesser technology than other technologies that are allowed to use the term by FDA, such as branched DNA.

This is a very important issue to Digene. They believe it would be unfair to the physicians and to the laboratories to deprive them of the information that this hybrid capture assay is signal amplified.

I am just going to quickly summarize. I cannot do it as well as the speakers before me, but I am going to summarize a few reasons that we are primarily relying upon, so that you can give your input to FDA on this very important issue.

First, the company believes that the peer-reviewed literature almost unanimously accepts the hybrid capture technology as signal amplified.

Digene has cited numerous articles to you. FDA hasn't cited one article which indicates, where there has been a review of this issue, that signal amplified is inappropriate.

You are all scientists. Typically, you know that scientific consensus drives definition. We believe here that the scientific consensus has been that hybrid capture systems should, in fact, be described as signal amplified.

It is not just the authors of these articles, it is the peer reviewers as well, who have all accepted signal amplification as an appropriate terminology for this system.

Moreover, secondly, we believe it is scientifically correct to label this product as signal amplified.

As discussed previously, the CMV assay alone has an amplification factor of over 3000-fold.

This is not a non-amplified detection system. It is not an insignificant multiplying labeled product. I have discussed this issue with Susan Alford(?) and

others in DCLD.

Their concern has been that the next guy in, who tries to use signal amplification for a simple ELISA test, is going to try to use a similar application in an improper way.

We are not going to presume to be able to tell FDA how each product that comes before them should be labeled.

We do know that, in connection with this product, signal amplification is the proper terminology, and signal enhancement is not appropriate, either scientifically or from a regulatory standpoint.

Thirdly -- and I know this is a matter of concern for every clinician -- is it clinically inappropriate for the clinician to be told that this is a signal amplification system.

I think the answer there is no. I believe they have shown you, in the slides which we just saw, that there is equivalent clinical performance to all the other available target and probe amplification detection systems.

Therefore, based upon the basic FDA principle and the clinician's principle of do no harm, you would do no harm by agreeing to describe this system as a signal

amplified system.

Therefore, in sum, we believe it would be contrary to the scientific consensus, as expressed in the peer-reviewed literature, to preclude the company from calling these products signal amplified.

It would be scientifically inaccurate to deny use of the term.

It would be misleading, and we believe it would be misbranding a product, actually, to call it signal enhanced. We believe the proper branding of the product is as a signal amplified product.

We believe it would place Digene at an unfair and unnecessary competitive advantage if they are unable to utilize this terminology.

I know we are not going for George Bush or Albert Gore here, but if you have to vote at the end of this session, FDA has given you an option, option C, which is nucleic acid, signal amplified solution hybridization assay.

We will accept that. We believe that is appropriate. We would hope that, when you are ready to vote, that is the direction you will take. Thank you.

DR. CHARACHE: I think at this time we can ask the panel if they have questions they would like to

direct at this point to the representatives who are here from Digene.

I would remind everyone that we should not be hearing new information, data, that has not been reviewed by the FDA. Questions from the panel?

DR. EDELSTEIN: I would like to direct a question to my colleague, Professor Kricka. This involves, in fact, what methods can be used to draw the line between a simple enzyme immunoassay and an amplified assay.

In your figures, you used multiples of enzymes that became orders of magnitude. Is there a way to, in fact, quantify this?

DR. KRICKA: I think it is easiest to look at this from the strategies that were employed in the assay design, how those individual strategies add up.

I think what you are trying to ask here is whether we can assign some numbers to these things. So, if you have 10 recognitions per 100 units of signal, and you have two labels for probe, that is worth another 100 that you multiply the first number by.

Then, if each labor is multiple signals, you then assign numbers to that, and go back into the equation, and then you work out a numerical score based

on calculations like that.

I am not sure anyone has ever tried to do that, to try to quantitate risk so that they can draw a line based on numerical scoring. That has not been done.

I think what you can do, you can look at the number of probes. You can quantify it by different quantitations.

You can count up the number of recognition events and, therefore, the number of probes. You can look at the number of labels. You can look at turnover numbers for enzymes to say what sort of signal you produce for individual labels within a given time. You can certainly do all of that.

I think the analysis that Dr. Lazar showed, where he went through the hybrid capture assay and tried to put numbers on each of the events, is the closest you are going to come to numbers.

I think this must be on an assay-by-assay basis where individuals who developed the assays and know what some of these numbers are, can assign the numbers and could provide a similar sort of analysis that Digene did for their assay.

DR. EDELSTEIN: Is there an objective way to do this, a non-theoretical way to do this? For example, you

heard the presentation that Digene has amplified 3,284-fold, was it? That is a theoretical calculation.

Is there a way to, in fact, establish an objective way to measure this?

DR. KRICKA: Ultimately, the answer to that is yes. You could set up, or attempt to set up, a simple one-signal, one-label, one-probe type of design and contrast that with the assay in question.

If you look generally at the immunoassay literature, and look at people's deliberations on sensitivity, and the arguments about what sensitivity means, there are very few examples where people attempted to do that with the assay, often, in fact, because the assays have been done by different people in different places, and no one has ever been in a position to pull off the shelf reagents and set up these sorts of assay designs and compare them one to one.

There are a few studies like that, but generally they are not. But in theory, you could do this.

Any manufacturer could scale back the assay design, back to one label, back to one recognition event, back to a label that gives the worst possible signal of all, zero.

So, it is possible to do this, but in fact,

apparently very little work like this has been done.

DR. EDELSTEIN: Finally, what is your practical suggestion to differentiate between a routine enzyme immunoassay and the amplification method.

This is a question, I think, primarily of precedent, and what is going to be presented to the FDA where there is a manufacturer who says, my enzyme immunoassay multiplies theoretically by a factor of 10.

What is going to be the practical way to sort this out, if any.

DR. KRICKA: I think the way to sort this out is to do the sort of feature analysis, and look at an assay and ask the question, do they have elements within the assay that would lead to amplification of the signal.

Do they have multiple recognition. Do they have multiple labels. Have they chosen labels which, themselves, will give multiple signals. Then, add those together and use that as a feature of the assay.

Obviously, then you might want to see some real numbers as to what the expected amplification is.

For this, I tried to contrast this to show that the Digene assay and also the trans-DNA assay represent, at the moment, the extreme of the spectrum.

Everything they can do to generate more signal,

by addressing each of the factors I have identified, by multiple recognition, multiple probes finding, multiple labels, multiple signals or individual labeling.

They have maximized each of those. So, they lie, very firmly, at one end of the spectrum. At the other end of the spectrum is the very simple assay where you might have one antibody, one label, one signal, less than one signal per radioactive isotopes, where you are waiting for it.

I think in defining the ends of the spectrum, that is fine. The middle ground, which will be your problem, I think, is a little more difficult.

I think you want to look for orders of magnitude, multiple orders of magnitude of improvement, which is what I think people think of when they think of the word amplification.

They think of something exponential, which is multiple orders of magnitude, not a simple, perhaps one order of magnitude of change, or a doubling or trebling of the signal.

DR. CHARACHE: Other questions? I have one question for Dr. Kricka. The 3,284 amplification enhancement, in comparing that number with the identification of 700 copies per milliliter, how do we

resolve that discrepancy?

Also, I am wondering about the 500 copies per milliliter, what matrix that was in.

DR. KRICKA: Let me deal with my part of the question first. Those were sensitivity detection figures, taken from an assay.

They didn't deal with the signal generation. This is the end phase of the assay. What you cannot distinguish in terms of the analytes in the reaction mixture, and includes factors which mitigate against sensitivity, such as non-specific binding and other factors which will reduce sensitivity in any assay.

Equating these things quite directly, I think your concern is that the 700 copies that we got, 3,000 was an amplification factor. We want to bring those together.

I think we are trying to compare things that are not directly comparable. You were able to get to 700 copies because you had a signal generation system which was amplifying, that allowed you to see those 700 copies above the background. I think that is the way to look at that.

If you have an amplification system, which is certainly more than 3,000 -- go 30,000 or 3 million --

then my expectation is that your 700 copies would drop proportionately, eventually reaching the background, below which you couldn't go.

DR. LAZAR: For CMV, most of the 700 copies is typically measured, first, by using plasma DNA and kind of a pre-matrix. Then we try to verify that in the actual clinical matrix.

It is quite difficult with CMV to actually get quantitative real clinical samples. There is no real validated way yet to quantitate absolutely the CMV numbers.

We used some quantitative CMV that we got cultured somewhere and someone counted it with an electron microscope.

We didn't think that was very accurate, but the numbers are similar when we dilute it into the clinical matrix, which is whole blood using CMV-infected cells in whole blood. We had similar results.

DR. CHARACHE: Another question?

I think we will move forward, then. The next presentation is from the FDA.

**AGENDA ITEM: FDA Presentation.**

DR. RAO: Good morning. My name is Prasad Rao. I am the lead reviewer for the device being discussed

today. It is the hybrid capture system CMV DNA Assay from Digene Corporation.

The 510(k) was cleared by the FDA in September of 1998.

The device is indicated for the qualitative detection of CMV DNA in human peripheral white blood cells in immunocompromised patients.

The issue today is not the performance of the assay, but the use of the term "signal amplification" as it relates to the device.

In February 1999, FDA received from the sponsor an amendment requesting the use of the term "signal amplification" in the product labeling.

In order for the FDA to perform its function and ensure truth in labeling, it is important for everyone concerned -- that means, the agency, the sponsors and the physician community -- to understand the signal amplification terminology and apply it correctly to device labeling.

I would like to present the FDA perspectives on this device labeling.

We are aware that the term "signal amplified" receives special reimbursement codes similar to nucleic acid amplification assays from the Health Care Financing

Administration, and other third party providers.

However, FDA decisions are independent of reimbursement issues. Our aim is to strive for appropriate labeling that is consistent with the safety and effectiveness of the device.

If the hybrid capture system is described as a signal amplification test, then many of the ELISA tests, numbered in the hundreds, may also qualify for the same classification.

The literature references are mixed pertaining to this device description. There are references in the literature where the Digene hybrid capture technology is mentioned without the descriptor signal amplification, and in some other papers it is classified as a detection test that does not involve any amplification.

It is agreed that multiple copies of the monoclonal antibody attach to the DAN:RNA hybrid, but does that qualify it as a signal amplification assay?

In an immunological reaction, where the antigen is a large molecule of 38,000 base pairs and has a molecular weight around 22 million daltons, it is to be expected that multiple antibodies attach to the hybrid.

However, the number of enzyme molecules that can be attached per unit length of the hybrid is much higher

in a classic signal amplification reaction than what we see in the hybrid capture system. I will come back to this point later.

Due to an oversight, FDA previously allowed digene's HPV assay to include the term signal amplification. The labeling change by Digene was proposed at the last minute, and was not noticed by the review staff.

The schematic diagram for the assay as provided by the sponsor is shown here.

In the assay method, specimen DNA is hybridized to specific RNA probes supplied in the assay kit. Keep in mind that there is no nucleic acid amplification involved here. The resulting DNA:RNA hybrids are captured onto solid surface by hybrid specific antibodies.

The immobilized DNA:RNA hybrids are reacted with alkaline phosphatase conjugated antibodies. The extent of the reaction is measured by chemiluminescence.

The question we have to address is whether the signal detected here is truly amplified.

I would like to point out that once the DNA:RNA hybrid is formed, in step two, it acts as an antigen in the remaining assay steps.

The hybrids are immobilized on the solid surface by antibody capture. Subsequent signal detection by DAN:RNA hybrid specific enzyme conjugated antibodies uses the familiar ELISA-type methodology.

For instance, ELISA assays for the direct detection of viral and bacterial antigens from clinical samples are known to follow similar signal detection steps.

As you can see, the issue of signal detection affects a range of tests. Therefore, we need to apply correct terminology for all such assays.

At this stage, we can ask the simple question, what is signal amplification. When compared to radioactive isotopes or fluorescein labels, the enzyme mediated detection systems are considered amplification reactions, because the enzyme does not just sit there on the target, but repeatedly catalyzes the hydrolysis of the chromoflor(?) and then the substrates.

That is the reason why ELISA tests are generally more sensitive than immunofluorescence assay tests. It is understood that the amplification is inherent to the ELISA technology and these assays are not explicitly referred to as signal amplification reactions.

Signal amplification technologies originated in

nucleic acid detection as alternatives to target amplification.

For a test to be labeled as signal amplification reaction, we would like to see "true signal amplification" to justify the special label.

One such reaction we can refer to for such attributes is the generally-accepted signal amplification reaction to the branched DNA assay.

I would like to remind you that CDRH has not cleared any devices containing branched DNA technology.

In a signal amplification assay as exemplified by the branched DNA assay, we see compound probes that have complementarity to the target and with the extenders that bind to amplification multimers.

There are multilevel probe hybridizations that amplify the capacity to bind multiple reporter molecules; that means the enzyme molecules. The branched DNA assay system is shown in the next chart.

In the classic nucleic acid signal amplification method, you see attempts to increase the primary signal by amplification multimers that bind to the extenders on the probes.

I would like to take a couple of moments to go over these various steps.

This slide is taken from AR quantiplex(?) HIV RNA, the assay is dated June 1996. In the schematic here, once the target is captured by the capture probes, then there are other set-up targets with extenders that bind to the target and then return to the extenders and bind to the pre-amplified markers(?).

The complementarity of each probe here, two such neighboring probes would attach to one pre-amplifier molecule. The pre-amplifier molecule binds to the amplifiers.

There is another level of probes that contain the enzyme, alkaline phosphatase, that attaches to them, and that would hybridize to the branches.

If we choose another technology, than the hybrid capture system that is under discussion, the 30 nucleotide long hybrid here could attach to one hybrid specific antibody.

Taking the other probes into account, you could have six enzyme molecules attach for the 60 nucleotide-long hybrid.

In this step here, each branch of the amplifier could attach three enzyme molecules, and then there are 15 branches and each amplify, eight amplifiers on each amplifiers, carrying possibly 60 enzyme molecules, thus

amplifying the number of enzymes you can attach.

Thus, the issue could be, six enzyme molecules attach using the hybrid system antibody, the hybrid capture system, to 60 enzyme molecules attached to a single amplification.

However, we are not using the branched DNA assay alone for signal amplification. In an immunological detection, the primary antigen antibody that may be detected by the use of enzyme conjugated secondary and tertiary antibodies that would form the so-called "Christmas tree" structures and, thus, increase the initial signal.

In comparison, hybrid capture system is a direct probe detection method where the primary signal is detected by enzyme-conjugated antibodies.

Granted, there are multiple antibodies attached to the hybrid. This situation is similar to biotinylated probe detection methods or other ELISA systems.

Functionally, the hybrid capture system CMV DNA assay sensitivity, as presented in the package insert, was comparable to the traditional CMV detection methods, such as antigenemia assay, shell vial assay, and cell culture.

As you have heard this morning, the sensitivity

of the assay could be different with different analytes and matrices.

FDA Issues. In the labeling of the device, the test is to be claimed as "nucleic acid signal amplified solution hybridization assay."

The review team has the following concerns about the use of the term "signal amplification" in connection with the assay.

One is, is the detection method used here similar or unique compared to other ELISA assays for antigen detection, or biotinylated probe detection methods?

Second, is there multiplication or replication of initial signal achieved in the Digene hybrid capture system which is expected in a true signal amplification reaction?

In order for the agency to make an informed and scientifically appropriate decision, we request the panel to provide advice and recommendations on the following issue.

Given the nature of the technology in this device and the performance likely to be seen by the use of this test, which of the following is an appropriate description of the Digene hybrid capture system assay:

- a. Nucleic acid solution hybridization assay;
- b. Nucleic acid signal enhanced solution hybridization assay;
- c. Nucleic acid signal amplified solution hybridization assay;
- d. Other recommended nomenclature.

The device was cleared as b, a nucleic acid signal enhanced solution hybridization assay. Thank you.

DR. CHARACHE: Any questions for Dr. Rao from the panel?

MR. REYNOLDS: Again, I note, it is somewhat apples and oranges, but how would the signal enhancement in the Digene assays compare to the signal enhancement of a branched DNA assay. Would you say it is comparable, more, less?

DR. RAO: I have noticed that in the branched DNA assays, they get detection by quantitative assay. Following that qualitative assay, it could be less.

DR. CHARACHE: Other questions for Dr. Rao? The manufacturer has requested five minutes added discussion at this time, after which we will have a brief break.

DR. LAZAR: Thank you. Thank you for this opportunity to speak again. I hope you all have a good picture of the bDNA, and I would request that it go back

up -- I see that the projector has been turned off.

One of the things that is a feature of the bDNA, is that those branches, where they bind to the target molecule, are very widely spaced.

The bDNA are not taking advantage of that space between its binding probes. That is where hybrid capture gets its amplification from, is by using every bit of sequence on the target molecule, to generate signal directly.

As the comparison to an ELISA test, yes, some features of the assay are similar to ELISA. But hybrid capture, as a technology, is making the antigen. In an ELISA assay, the assay does not make the antigen.

The assay design controls the size of the antigen. It can be bigger or it can be smaller. That is not what happens in the ELISA assay. You have the antigen that is in your sample.

There is no RNA:DNA hybrid in the sample that we are detecting.

Furthermore, I believe that if you actually do the calculations, that the number of alkaline phosphatase molecules that you end up with, with bDNA and hybrid capture are very similar.

In fact, in an 1992 article by Wolfpach(?),

which I think is in a reference package we sent you, it specifically states that a bDNA assay bound 400 alkaline phosphatase molecules.

We are certainly far ahead of that now, and at that time it was described as signal amplification.

bDNA has probably also advanced beyond 400 alkaline phosphatase molecules, no doubt. But still, we are both orders of magnitude ahead of a single probe, single label scenario.

I would just like to clarify a little bit about that HPV test. I know that the DCLD reviewers mentioned that, and we would just like to take exception and say that the labeling change was discussed explicitly in the introduction to this submission, and was not just slipped in at the last minute.

There were several pages of discussion of that labeling change in our first submission, or that supplement.

Finally, I would just like to underscore that, from a sensitivity standpoint, I know that we can't present data on it today, but we have presented data in the past showing detection down to 100 copies for a similar analyte, to bDNAs 50 copy level.

So, it is the actual sensitivity level. Fifty

and 100, as you all know today, is really meaningless; it is not significant.

I think in any target that we would choose to compare directly with bDNA, we are always going to be very similar in sensitivity.

I ask you to take these things into consideration in your deliberations today, and thank you very much for your time.

DR. CHARACHE: Thank you very much. We are running a bit behind, so we will take a break until 11:30, and then reconvene for discussion, and you can ask questions subsequently to any of our previous speakers.

[Brief recess.]

DR. CHARACHE: Our next order of business is to ask whether there is anyone from the public, any attendees, who would like to speak.

I don't see anyone. We have not been informed ahead of time that there would be a public speaker. So, we move on, then, to the panel discussion.

**AGENDA ITEM: Panel Discussion and Recommendations.**

DR. CHARACHE: Reminding the panel, that this is what we have to consider: What title would be the best, most applicable for this test.

It is a matter of linguistics. We have heard that it has major implications in terms of some aspects.

What is most appropriate to the technology and the information that we have received.

Should this be considered to be a nucleic acid solution hybridization assay, nucleic acid signal enhanced solution hybridization assay, nucleic acid signal amplified solution hybridization assay, or other nomenclature which the panel would like to propose.

I think we have heard some discussion of the advantages of each a, b and c, and some reservations. I would like to begin by asking someone from the panel to express their views, which we can build upon, and if we have questions of our previous speakers, we will address them at that time. Who would like to start?

DR. SPECTER: I will start off by saying that my view is that a, b and c are all acceptable, because I don't see that they are necessarily exclusive; rather, that they are inclusive.

I based that on four things, and I will finish with one question.

First off, Dr. Rao indicated, when he spoke, that ELISA is, in fact, an amplification technique. Therefore, by nature, this would have to be.

He also told us that we don't use the description term, amplification, for ELISAs. That is probably why this issue is here in the first place.

Secondly, most simplistically, the words enhancement and amplification are synonyms, as best I understand them. So, it seems to be a semantic question, unless common usage is important, which I believe it is.

We have heard a number of times that signal amplification is already in common usage for this technology in the literature.

Therefore, it would not be confusing to continue to use such a term. That would make c acceptable, and it would also make b acceptable, since they are both synonymous.

Finally, the thing that I think is most terribly important here is this question of safety and effectiveness.

I honestly believe that, regardless of which of these three terms is used, there is no compromise of either the safety or effectiveness of this particular assay.

The important thing here -- and some of us were discussing this -- is that really, for the most part, this is a laboratory issue.

Clinicians seldom walk into the lab and ask you for a particular type of assay as opposed to a test for a particular analyte.

I think the important thing is, they want an answer for a clinical condition, and they are relying upon the laboratory for that technology.

To me, the laboratory is very comfortable with signal amplification related to this, and the physician probably doesn't really know what you are talking about, number one, for the most part, and probably cares even less, as long as he gets the answer he needs.

That brings us really to the final question that may impact on any of this, and that is the distinction of whether ELISA is an amplification assay for public documentation, and whether we need to be addressing that.

I think that really is a somewhat separate issue, as this is a special technology, it is not straightforward ELISA, it is more complex than a simple ELISA.

I am actually in favor of using any of these three. If it is a question of whether ELISAs should be allowed to be discussed as amplification techniques, I think that is a separate, but probably a very important issue.

DR. CHARACHE: Other thoughts?

DR. HAMMERSCHLAG: I think it goes beyond semantics. I think there is a potential here for a lot of confusions, especially with ELISAs.

Just using, let's say, chlamydia testing as an example, there is a world of difference between the performance of ELISA and the amplification using nucleic acid amplification tests that are out there at this point.

You are talking about wide differences in being able to detect the organism. The performance parameters are so much different.

I think that, not only do the physicians frequently not care, but I think that they are getting totally confused about the technology.

If they are confused about the technology, they frequently don't have much choice about the technology, and those decisions are made by whatever laboratory that is contracted by their HMO for whatever hospital where these specimens are sent out. I don't know how many of these tests would actually be run in-house, but that is another issue to be discussed.

The question is, using the term amplified, do we begin to lump this particular test into the same

category, for instance, as a nucleic acid amplification test, with those performance parameters.

To me, I think that is probably one of the major issues, rather than calling it enhanced. Practically, these tests should always be evaluated on performance, not so much on whether it is amplified or how they title it, because technologies could be different.

I think there could be a problem when ELISAs come up, and whether they truly are performed as nucleic acid amplification tests.

DR. SPECTER: If I could just reply back, I was not advocating that we want to refer to ELISAs as amplification technology at all.

I was making the point that a critical issue, as you addressed, is the safety and effectiveness of the testing. I have no perception that this affects that.

DR. GATES: I think part of the issue, too, is that we are making a discussion based on label claims, and we are talking about signal amplification as an implied or informal claim.

I think that if we are going to be locked in because of that, we can go either way on it. I think the issue is that that is the wrong thing to focus on.

Do we want to have some sort of definition as to

what the test is, I think the labeling makes that explicit and we don't need to rely on signal amplification and establish some kind of informal definition. I think you can address by explaining explicitly about what the test does.

DR. CHARACHE: I will ask everyone to give their names for the recording.

I am wondering -- and I am not sure who would answer this -- what would prevent a manufacturer of an ELISA assay from insisting that his be considered signal amplified. Where is the break in this type of continuum?

Would this present a legal issue? If so, is it one that should concern us? Dr. Gutman, can you help us?

DR. GUTMAN: I guess the only one that might present an obstacle in this labeling would be the FDA, who might not be entirely pleased with the broad use of that labeling in that context.

Of course, we are here to listen to you, so if you suggest that is a good idea, we will consider it, if you think it is a bad idea, we will probably continue our current practice.

I think the issue in the division is whether anyone has yet to challenge us with that request.

DR. SPECTER: I raised the issue before, and

that is why I made a distinct point about it at the end of my comments.

I really think that it is an issue and I think it should be addressed before, rather than after, there is any desire to do this, and the definition would be created as to what can be referred to as an amplification assay, either to control ELISAs using this, or to give guidance prior to someone trying to use it. It would be very helpful.

DR. GUTMAN: Again, it is a matter of history. No one has come forth with an ELISA and made the claim that it is an amplified technology.

What our starting point is, when we first began reviewing and thinking about the concept of amplified as a problem was in the context of nucleic acid amplification techniques.

So, that is the starting point at which the term appears. Our angst over this is the issue that the reviewer raised, and I think that you are also raising, and maybe it doesn't have an answer.

It is the issue of, since we have gotten into the technique -- we have not considered and equivalent, and maybe that is the wrong call.

The concern is on the table. The concern as it

percolates back, this is a snazzy technology. I think that, no matter what we do, this is a snazzy technology.

So, the deal here is, is it snazzy enough, or is there some semantic way that we bring this in, or is there some alternative term that allows it to be snazzy without necessarily confusion with broader technology?

Do we not worry about this, and bring in some English professors and let them worry about it, let HCFA worry about it, let the marketplace worry about prejudices or lack of prejudices, and move on to the next submission.

DR. SANDERS: This goes back to the question earlier, how much amplification is amplification. I am not clear that we have any answers to that question.

MR. KAHN: I hesitate to interrupt, but I just want to make one point. That is, that I think Steve is absolutely right, that the way that this would work in the future is not quite as difficult as the panel might think.

Typically, the way that FDA handles these situations is that, when an issue like this comes up, there is a consensus reached and there is a guidance document issued by FDA, based upon panel input and expertise within the agency.

Then they advise everyone, ELISA manufacturers and manufacturers of snazzy tests, what is appropriate and what is inappropriate.

Here, I believe Dr. Kricka gave you a fairly good basis that a guidance document could be based on, which would be multiple labels, multiple recognition events, and increased by orders of magnitude.

There are lots of ways that, in a guidance document, FDA can advise the industry as to what is appropriate, so that the office of compliance within CDRH doesn't have to go out every day and tell ELISA manufacturers, you shouldn't be calling your product signal amplified.

I think in terms of how FDA typically works, this is not that difficult a problem. Thank you.

DR. CHARACHE: Obviously, the FDA is concerned, because they brought this issue before the panel to consider further.

DR. EDELSTEIN: I think there are two issues, as Steve has pointed out. One is the specific issue of whether this product merits the label, amplified assay.

The second is whether and how the panel should provide guidance to the FDA on the use of this term in the future, in a more global sense.

My suggestion would be that we first decide on the labeling that specifically addresses this question and then, secondly, if the committee would allow it, is to have a general discussion of how we might approach the definition of this term, in terms of giving guidance to the agency.

DR. CHARACHE: I think I am hearing -- and I will ask for comments -- that one of the issues here is the linking of the word signal amplified versus amplification technology in general.

I know that the concept of amplification technology gets a little confused between the specific application, and PCR or whatever, that has a different connotation and implication in the minds of the clinical users and in the minds of the laboratory scientists as well.

The word amplified, in terms of microbial products gets very, very murky when you try to make sure that the word amplification remains attached to the word signal.

Then you have to decide whether the signal that is being amplified is a microbial product, or the chemiluminescence.

I think maybe some of these hook-ups, which are

used in very muddy ways now in the literature, and in this discussion, are what the FDA is struggling to address. Dr. Gutman is nodding.

I think one of the things the panel should be thinking about in terms of what the modifier of the word signal should be, is how to make sure that the user isn't confused in this very murky way.

I think that the problem with the signal amplifier is this separation, and that there is an impression that it has a power that it may or may not have.

The point was made by a number of people that that power is a function of the assay and the matrix in which the events take place, as well as all the rest of it. Of course, this is why there is such a disparity between the theoretical number of copies you could detect and the actual number of copies, which clearly carries over into many disciplines.

If we could show less than one copy by PCR of CMV, that is not clinically relevant. It just happens to be the experimental definition.

Can we think about that issue, of how to use a phrase that is fair and appropriately describes the unique technology, but which doesn't get us into this

confusion between what is amplified.

DR. RELLER: It seems to me, as long as you are dealing with enhanced and amplified and solutions, there is no way to get around the ambiguities inherent.

When one considers all the possible uses for the emphasis on this exact wording, one could envision, for d, snazzy, enhanced, fourth-generation nucleic advanced signal amplification solution hybridization assay.

One could simplify that to simply snazzy assay for CMV or SNAC.

DR. CHARACHE: It is almost that.

DR. RELLER: D and C, and what I propose for D, are ugly germanics backing up adjectives, that lead to obfuscation.

Rather than getting into an ever-deeper morass, what about simply calling this a nucleic acid hybridization assay for CMV.

These middle steps -- there are a lot of ways to do it. What seems to me to be the heart of this assay is a hybridization of DNA and RNA that is long, that enables a lot of antibody to be attached.

It is connected with an enzyme that enables a bright light, and it works.

I don't think we are going to get out of this

by, you know, dealing with terms that have so much ambiguity, as opposed to simplifying it and then let the test speak for itself, and the clinical applicabilities thereof.

DR. TUAZON: Can I just make a comment as a clinician? I have to echo the comments by the panel members, that the bottom line here is the performance of the clinically, the sensitivity and the specificity, and the efficacy, as well as the diagnostic usefulness. To us, it doesn't really matter how we label this.

When we go to the lab we say, you know, we need the most sensitive and the most specific study for the diagnosis we are going to make.

DR. EDELSTEIN: I would like to counter Dr. Reller's comments. While it is true that a more specific definition would be acceptable to us as both clinicians and laboratorians, the problem is that we need to have a level playing field.

The playing field is not what is going on in the laboratory. It is what is going on in marketing and reimbursement.

If we don't allow the use of the term amplification, that is going to create an uneven playing field in terms of reimbursement and being able to market

the product.

Personally, I am not certain that this is something that we should be discussing, as a committee that, overall, decides on issues of safety and effectiveness, as opposed to making a decision regarding labeling in terms of marketing and reimbursement.

That is what is on our plate. I think that we have to keep that in mind.

DR. CHARACHE: I certainly think that we understand that that is why the issue is before the panel. I think there are very major issues for the manufacturer.

DR. WEINSTEIN: I guess part of my problem here is this issue that Paul and Elsie were trying to get to earlier, before the break.

We are dealing in partly semantics based on orders of magnitude. The question is, what do most laboratorians and clinicians think of when the words, nucleic acid amplification, are all in the same phrase.

I think that most people think of orders of magnitude in the millions or billions of copies, as happens in a PCR kind of assay.

So, it is semantics, but it is also orders of magnitude. I think that is where the dilemma is.

DR. CHARACHE: I would like to see if we can suggest another D in addition to SNAC. That is, can we avoid some of the ambiguity by using the same words, but in a different order?

The word signal kind of gets lost here. You are not sure whether the signal is amplified because the nucleic acid is amplified. That is causing confusion.

Could we turn it around by simply calling it a signal amplified solution, nucleic hybridization assay?

The words are the same, the meaning is the same, but you clarify it by making it clear what is amplified.

This also applies to the branched DNA, and probably should. Comments?

DR. WEINSTEIN: I guess if you would do that, you could even omit the word solution. Could you call it signal amplified nucleic acid hybridization assay?

DR. CHARACHE: That would certainly further clarify it. Certainly, all of us know that, with hybridization assays, it doesn't really matter very much, whether it is in a solution or not, as long as they do the job.

MR. KAHN: You could throw in some commas or hyphens in various places, signal-amplified.

DR. CHARACHE: That would better define it.

DR. RELLER: The sequence, and the best English for this, and actually, when I made my comments about the stacking, I purposely added a few more to get my point across earlier.

At its heart, this is a nucleic acid hybridization assay with signal amplification, or you could say nucleic hybridization assay with signal amplification.

Whatever you want to do, I think where there may be some comment agreement, to get out of the dilemma, I is to put the emphasis on the multiplication on the signal as opposed to the multiplication of the extant DNA in the sample that is somehow delivered, extracted, either added to a number of white cells, et cetera, in the first place.

I frankly would favor scrapping the irrelevant solution, simplifying it and saying, a nucleic acid hybridization with signal amplification. Then it gives one the opportunity to be more precise.

Other descriptions of assays, with whatever the component is, put the principal emphasis on the multiplier.

DR. WILSON: I think that the dilemma here is that, to some extent, we are trying to pigeonhole

emerging technology with existing definitions.

I think this is an important issue because of the precedent that we need to think about. As new technologies come out, we will always grapple with trying to use descriptors and apply different technologies.

I agree with what Dr. Charache and Dr. Reller said. What we need is a very precise way to describe what, in fact, the assays actually are and what they actually do, rather than trying to lump them into another category.

I think that is the best way to avoid future problems such as this.

DR. EDELSTEIN: Since this panel is providing advice to FDA, my suggestion, it seems as if we all agree that the term amplification should be, or can be, included as a descriptor of the test.

We should let the company and the FDA come to the terms of how the English will be put together, whether it will have active or passive voice, hyphens, commas, or in what order you put the adjectives.

DR. SPECTER: I just want to follow up on Paul's comments, because I agree with him completely. I think there are two major issues.

One is the issue of whether signal amplification

could be used or not. The other is that there is clarity in describing what is here.

Therefore, I believe we should be recommending whether or not signal amplification should be used -- and I a lot of people feel it should be -- and the fact that this is a nucleic acid hybridization assay.

I think those are the two points we need to make a recommendation on, one, whether we can use signal amplification and, two, whether we link nucleic acid hybridization and let the final terminology be worked out by FDA and the company.

DR. CHARACHE: I think we are agreed that we have to make clear that it is the signal that is being amplified.

We have suggested two Ds, which are ways of clarifying that it is the signal that is amplified. One would be signal amplified nucleic acid hybridization assay, and the other would be nucleic acid hybridization assay with signal amplification.

These are two suggestions which would make it clear that it is the signal that is being amplified and not the nucleic acid that is being amplified.

Yet, it emphasizes the fact that the signal is being enhanced.

Can we get a sense of the panel in terms of the -- let's take a look at A, B, C and D. Let's get a sense first, whether the panel feels that it is fair to the manufacturer and is a good descriptor compared to the other options we have heard, to use solution A. I don't know, just go around very quickly and say yes or no. If you say yes, amplify it, and if you say no, we will go on.

DR. HAMMERSCHLAG: From the discussion I would say no for A.

DR. SANDERS: I prefer no for A.

DR. WEINSTEIN: No.

MR. REYNOLDS: No.

DR. GATES: No.

DR. RELLER: No.

DR. SPECTER: No.

DR. EDELSTEIN: No.

DR. TUAZON: No.

DR. WILSON: No.

DR. CHARACHE: A is gone from our recommendations. What about B, which the FDA added to help avoid the word amplified?

Does the group feel at this point that that is a necessary substitution, given the manufacturer's

presentation?

DR. HAMMERSCHLAG: I think, considering that we have another alternative, I am going to say no to that as well.

DR. SANDERS: No.

DR. WEINSTEIN: I have mixed feelings. It would certainly not bother me to use that term, but I think the other terms that have come up in category D are probably preferable.

MR. REYNOLDS: No.

DR. GATES: No.

DR. RELLER: No.

DR. SPECTER: No.

DR. EDELSTEIN: no.

DR. TUAZON: No.

DR. WILSON: No. I think the issue here is that we can't define amplified. How can we define enhanced?

HAMMERSCHLAG: It is unusual to use a Thesaurus, but I think scientifically for a given -- why don't we just move on to D because I think C is --

DR. CHARACHE: Could we have a show of hands on those who prefer to stick with C at this point, having heard the discussion and the debate? Could you start?

DR. EDELSTEIN: Yes. I think the language to me

is clear. It says what is being amplified. I don't have strong feelings about it. I think it is important to include the words signal amplified in whatever label is used.

I feel a little bit uneasy in suggesting alternatives. I am not sure that is really what we should be doing.

What I am saying is that I think that is just as fine as any other term, or any other descriptor which uses signal amplifier.

DR. CHARACHE: So, you continue to concur that there should be some emphasis on the signal amplification, but you don't really care whether it is C or D?

DR. EDELSTEIN: I haven't heard a definitive D. So, I need to hear what D is exactly, but I assume that D includes something with signal amplified.

DR. CHARACHE: Let's consider the two Ds that have been named. One is signal amplified nucleic hybridization assay. The other is nucleic acid hybridization assay with signal amplification.

Can we perhaps see if the group would agree with D or come up with any other D, and then go to the recommendation that was made, and let the FDA and the

manufacturer decide which would be most appropriate for their purposes.

Are we ready to do that yet?

DR. RELLER: Just to enhance things, speed things along, I would like to propose a vote. If the vote fails, obviously the sense of what we want -- or I shouldn't say want -- but suggest is with the agency.

Either a specifically D passes or the language falls to the agency.

DR. CHARACHE: All right, I think we can pursue that course. There are differences between the two statements.

I think the general consensus is that we would prefer some of the Ds to the C, because of the ambiguity we find.

The second of the two D recommendations, nucleic acid hybridization assay with signal amplification, emphasizes that this is basically a hybridization assay.

The first doesn't make that emphasis.

Now, Barth, do you see it that way?

DR. BARTH: I see it that way, but also, it is better syntax and it sounds better and it is said more readily.

As a consequence of its clarity, it ends up

being a better description of what we are talking about.

Consequently, I would suggest that if we vote on it, and if it doesn't pass, that is the agency's task.

DR. CHARACHE: Recognizing that we are an advisory body and not a decision-making body. Other thoughts?

I am not sure if we have to decide which one of these to vote on first.

MS. POOLE: Just to remind you that we are giving recommendations. We are not actually voting. We are just accepting your recommendation.

DR. CHARACHE: We are voting on what recommendation.

MS. POOLE: Yes, in essence, it is not really voting, and not voting approval or not approval. We are just taking recommendations for which is preferable, a or b of these.

DR. CHARACHE: I think the group has recommended that it not be A, B or C.

DR. SANDERS: That is what I wanted to clarify. Within D, we are about to determine whether or not to recommend, not hard but soft, or soft, not hard.

DR. CHARACHE: I think which is more grammatically correct.

DR. SANDERS: Then whether or not we would recommend D or C, or has C been totally discounted?

DR. CHARACHE: Perhaps we should discuss C again, bearing in mind Paul's comment.

DR. SPECTER: I think it would be easiest if we just went around the table and discussed whether those option were acceptable to the panel members, and take a vote on acceptability.

Then we will know if we can turn it over to FDA and the company to work out, or whether none of them is acceptable.

DR. CHARACHE: I think that is a wonderful suggestion. So, what we would do is ask about the acceptability of the Ds. If anyone has comments on the acceptability that they would like to add, I think we would like to hear it at this time.

DR. HAMMERSCHLAG: I think either D. I am not a grammatical expert at this point. I think it describes what is actually going on with the test, and people will know what is being amplified.

The agency will probably have to go into some discussion or definition of what amplification means, because I see this coming up again an again.

I would leave which version of D to be selected

up to the pundits who are the syntax experts. At this point, it has been a long time since I have done that.

DR. SANDERS: Leave it up to the experts is what my opinion is, although I think it should be signal amplified.

DR. WEINSTEIN: D-1 and D-2 are both acceptable.

MR. REYNOLDS: Both the Ds are acceptable. My only caveat is that C is currently used for other similar technologies and you might have to look at changing the labeling on those packages.

DR. GATES: I have no problem with C. My preference is D-1 because it is a little more economical and has fewer commas.

DR. RELLER: I don't think there are any commas in either of them. I think I have made my points already.

DR. SPECTER: I feel both are acceptable.

DR. EDELSTEIN: I think C, D-1 and D-2 are acceptable.

DR. TUAZON: I think both options are acceptable.

DR. WILSON: D options are both acceptable, prefer D-2.

DR. CHARACHE: Thank you. I think we have

provided our basic recommendations. Are there any other suggestions or comments that you would like to offer at this time?

Hearing none, we would hope that we have been helpful. I think we can adjourn at this point. We will reassemble at 1:30 promptly.

DR. GUTMAN: I was wondering if any hearty souls here would actually be willing to help us identify what we should be interpreting as signal amplification.

DR. CHARACHE: Anyone who is willing to assist with this, would you please contact Steve Gutman, who would appreciate your volunteering.

DR. GUTMAN: Anyone who would want to extemporaneously do this.

DR. CHARACHE: Right now. All right, what would be a reasonable way of saying that something is amplified because it is expected to be, to that degree, and what represents a break-through technology which warrants this terminology of signal amplification. Any thoughts?

I will start, just so people can shoot. I think we would say that ordinary technology in which there is an amplification due to a detection system only, as opposed to amplifying the signal that goes to the

detection system, would not be considered signal amplification.

A new detection system would not, in itself, represent signal amplification.

DR. WEINSTEIN: I was sort of thinking of it in a different way. If you want to use the term nucleic acid amplification, it seems to me that you are talking about magnitudes of millions of copies as opposed to a few thousand, which might have a different definition.

DR. CHARACHE: I think now we are talking about just amplifying the signal, as in the case of the branched DNA or the Digene technology.

Any other thoughts as to, at what point it becomes a signal amplification?

DR. RELLER: Isn't there -- as some point, isn't there a role, Steve, for the concept of comparatives? If one looks at the process and where the principal multiplying step is, that if one has comparable sensitivity or equity in labeling and evaluations, that if one has a product like this one, that is with signal amplification in the description, that if something else has a signal that is multiplied, that is of comparable sensitivity to something that already exists, that it would get a comparable label.

If it didn't match that performance, it wouldn't. If the performance greatly exceeded, we will be back in a couple of years about what the other adjective is.

In the end, there may be those that -- that if one has two diagnostic products with amplified signals, that one of them has better, for one reason or another, performance than the other one, but as long as one does as well as penicillin for streptococcal strep throat, that that concept -- that it would get a comparable label.

One has to start someplace, and when everything gets outmoded, then you have a different set of comparatives.

DR. GUTMAN: This is a tough question and I don't want to beat a dead panel into the ground. If folks do have thoughts, we do have the opportunity -- your discussion has been helpful.

I assure you that it is not a usual safety and effectiveness issue. It was a passionately important issue to us, and I think the sponsor as well, or we wouldn't have both agreed to bring this before you and spend the time.

I do appreciate your time. In my view, it is

more than just a semantic issue.

If you do have thoughts on the definition or how to develop guidance or how to developing labeling, we would certainly ask anybody, including the sponsor or the people in the audience, and certainly the panel, to share their thoughts with us.

Although I think you have been helpful, I think there is still some potential future ambiguity there that we would like help on.

DR. RELLER: Along those lines, if one doesn't have a basic comparison of performance, as opposed to getting into arguing about the theoretical number, whether you multiply all these things together, then I think one gets into a situation where the end number being 2,782 and 2,793 and you know, it is tough.

DR. GUTMAN: Or when you look at performance you get different designs and different populations and it can get quite tricky looking at variable performance.

We also don't have very crisp performance standards for any technology like this. So, although it resonates personally with me, I think we should try to address it.

At least what I have heard is that we need to work toward precision in the labeling so that it is

clear, whatever is going on, we are as close to the truth about the technology as we can be. That is one of the things I carry away from this. Thank you.

DR. CHARACHE: We will re-adjourn.

[Whereupon, at 12:29 p.m., the meeting was recessed, to reconvene at 1:30 p.m., that same day.]

A F T E R N O O N    S E S S I O N (1:35 p.m.)

DR. CHARACHE: We are going to get started. We are reconvened to discuss premarket approval application from Gen-Probe, called Gen-Probe Amplified Mycobacterium Tuberculosis Direct (MTD) Test.

The test is a target amplified nucleic acid probe test used for the detection of Mycobacterium tuberculosis complex in sediments prepared from sputum, induced or expectorated, bronchial specimens, or tracheal aspirates from patients with smear-positive respiratory specimens.

The device indications were modified to include AFB smear-negative respiratory specimens with a diagnosis of active pulmonary tuberculosis disease.

It has been approved for smear-positive respiratory specimens. The issue that the panel will address this afternoon is whether we wish to suggest that smear-negative respiratory specimens may also be tested by this method, to establish the diagnosis of pulmonary tuberculosis disease.

We will begin with the sponsor presentation, Glen Frieberg.

**AGENDA ITEM: Premarket Approval Application.  
Sponsor Presentation.**

MR. FRIEBERG: Thank you very much. I am Glen Frieberg from Gen-Probe. I would like to begin by thanking our panel members.

We want to thank you for your review, but we would also like to thank you for your time. We know it takes a lot of effort to come here today and do this.

I would also like to thank the FDA review team. A lot of time has gone into their review. They have been very open, and communicated with us continuously through the process and we appreciate that.

Consultants for Gen-Probe today, Dr. Catanzaro and Dr. Woods, are both being reimbursed by Gen-Probe. According to the rules, I would like to tell you that we are not a public company, so they have no interest in our company.

Dr. Catanzaro will be providing our last presentation, after which I will provide a brief summary.

Dr. Woods will also be here, if there are any laboratory questions toward the end of the presentation.

This is the group we have brought from Gen-Probe. Not everyone is presenting. I will go through how we are going to try to organize the afternoon for you.

At the conclusion of my introduction, Vivian

Jonas will summarize our experience with the marketed product.

I would like to clarify one thing that was said in the introduction, and that is that we are not here for a PMA approval. The product is already on the market. We are here with a PMA supplement.

DR. CHARACHE: I apologize.

MR. FRIEBERG: I will be showing you the enhanced intended use. When Vivian Jonas does her presentation following mine, she will be talking about, as I said, the intended use of the product, the intended use, and review selected sections from the package insert in response to feedback from the FDA and what we have received from the panel thus far.

The goal of our vote this afternoon is in regard to our request for five words, "and negative," "or negative," and "either." That is what we are trying to change in the package insert.

As a marketed product, you are probably all aware that these additional uses could be brought in under CLIA, perhaps, but we think the right thing to do is get the labeling updated and add tables that the panel and FDA feel are appropriate to give the end user the information they need to use the product safely and

effectively.

Now, some of the proposals that we have seen, and that will be discussed by the FDA, moves around more of the proposed intended use.

I would like to reiterate that we would like to stick with just what we have added with the five words, because the smear-positive approval is already out on the market. Any other changes could affect the prior approval and we do not desire that.

One other item I would like to remind the panel on, and to reiterate, that the FDA regulates products and their labeling, not the practice of medicine.

Sometimes, in situations like this, we drift a little bit into how we practice, rather than the safety and efficacy of the product.

I have two slides on the importance of what we are trying to do today. The first is the importance of a smear, and then indications for use.

The bottom line is, in the trial that you all reviewed, we were able to identify patients with tuberculosis that would have been missed, had the product been restricted to smear positive. We believe that is important.

As you know, diagnostics are also used for rule

outs. We want to provide this expanded intended use so that the clinicians have the ability to have another adjunctive diagnostic to be used in making the clinical diagnosis of TB. That is really all it is, as an adjunctive.

The MTD~~TB~~ test, to the best of my knowledge, and all the information that I have been able to received, is that it won't be used as a stand-alone. That is just not the way laboratories practice.

We will now proceed with Vivian Jonas' presentation. We look forward to discussing her submission shortly, after Dr. Catanzaro completes his presentation. We will do our best to stay on track, and to do so, we would request that, if possible, hold your questions until the end. Thank you.

MS. JONAS: Good afternoon, everybody. I am Vivian Jonas and my group at Gen-Probe is the group responsible for developing this product in R&D.

This afternoon I would like to talk to you about one issue that the FDA asked that we addressed, and that is the differences between MTD and -- and I am sorry we used the word enhanced. It was a marketing thought.

The word was used to demonstrate the difference between the first MTD test which was approved in 1995,

and the second one which was approved in 1998. So, I apologize.

I want to give you a little bit of information on how MTD has performed since the launch of MTD-2, and some of the questions addressed to Gen-Probe and to you all by FDA on the package insert, and then some conclusions.

The first thing is the differences between MTD, which was approved on December 15, 1995, and the second version of the test, which was approved in May of 1998.

We increased the sample volume nine-fold in an attempt to increase sensitivity. That required a decrease in the specimen dilution buffer volume.

The volume of lysate going into the amplification was apparently decreased from 50 microliters to 25 microliters. However, there is really more than twice the amount of sample going into the actual amplification reaction.

The amp time was decreased from two hours to 30 minutes. The selection time was increased from 10 minutes to 15 minutes.

I am sure you are all aware that there was some trouble with a perceived cross reaction with Mycobacterium ~~canzatsii~~ kansasii in the field.

When that surfaced, we worked with a laboratory that identified it, as well as our clinical trial sites, and the 15 minutes was adopted to eliminate that cross reaction.

The total time to result has been decreased from five hours to two-and-a-half to three-and-a-half hours. So, this is truly a one-day test.

Changing topic a little bit, we would like to go over the complaints that we have received since June of 1998 when the product was launched, until April 21 of this year, when I have this analysis done.

There have been seven complaints in that period of time, which is almost a year. There was one broken bottle, two shipping errors, three contamination events which our technical service department determined was procedural error and was successful in taking care of, and one potential inhibition.

That complaint rate, since the launch of MTD-2, was .0002 percent, based on complaints per total test, three zeroes. I always do this; I always put two in.

The entity test is quite robust out in the field, in everyday usage. We haven't observed any performance issues.

Now, there are many customers, as you well know,

who have validated MTD according to CLIA, for use with smear-negative samples. We have not heard of any issues of safety, effectiveness ~~inf~~ those customers' hands.

In addition, MTD has been available in Europe since 1992 and in Japan since 1994, with no restriction for smear.

Now I would like to turn to some questions that the FDA had in terms of the package insert.

One had to do with whether the warnings currently in the package insert were sufficient to guide the laboratory into using the test.

We want to make it clear that nobody in this room, or in Gen-Probe or anywhere is suggesting that we stop doing culture.

It is imperative that we do culture, not only to define mycobacterium other than tuberculosis, but also to address susceptibility.

In fact, MTD is a better test to help the physician make the diagnosis. It is not a stand alone. I think the warning -- which is the first bullet there and I am not going to read it -- is sufficient to do so.

The other question has to do with, are we being sufficiently clear on restrictions in terms of who the test should be used, and when should it be the intended

use, which is inpatients suspected of having tuberculosis.

There is an additional warning that says that the test is not being used to diagnose patients or follow patients through therapy. So, we are not looking for a claim to do that.

Do we have sufficient information in the package insert overall? Well, there is standard data including sensitivity and specificity, and we have broken that out in the performance of the test from patients with smear-positive as well as smear-negative specimens.

Every laboratorian can look at the package insert and determine how the test performs in those populations of specimens or patients.

There are data showing MTD performance with respect to increasing numbers of specimens. FDA has spent a lot of time asking us how the test performs with increasing numbers of specimens.

We don't believe we should tell the customer how to practice medicine, but we can give him the information we have on how the test performs whether you have one, two or more specimens.

Another question is, can a negative MTD result be interpreted without inhibition testing. The

unequivocal answer, in our minds, is yes.

In this particular study, we had 71 patients diagnosed with TB. There were 10 patients that were MTD negative and nine were tested for inhibition.

For two patients, all specimens were inhibitory, and one of those patients only contributed a single specimen.

That should be sufficient information based also on the fact that the negative predictive value of MTD for smear-negative specimens or patients was 96.4 percent, which is quite high, and culture, in fact, is not any different.

If you have more questions on inhibition after Dr. Woods or Dr. Catanzaro, I would be happy to answer your questions.

What is the appropriate interpretation of an MTD-negative result for smear-negative specimens? This is taken directly out of the package insert.

We didn't pick up M. tb MTD ribosomal RNA. That could have been caused by a variety of things. Either the person doesn't have TB, there might be lower numbers of TB in the presence or absence of mycobacterium other than tuberculosis, or there might be specimen inhibition.

Just like any other test, if you don't believe

the result and the physician really thinks that the person has tuberculosis, get another specimen and test it. That is exactly what we have.

Can a single positive MTD result, using the criteria and the directions for use, be considered definitive evidence for MTD in all patient specimens?

It is a clinical diagnostic. The physician needs to diagnose or determine whether or not the patient has TB. The test can't do it.

All we can do is give positive predictive values and show, as you see the data up there, that in a smear-positive patient, the positive predictive value is 100 percent. In a smear-negative, it was 75 percent.

I would like to conclude by saying that the current package insert allows for the safe and effective use of MTD for its current claims.

Addition of clinical data with respect to smear-specific information is all that is required for this application.

MR. FRIEBERG: Thank you. Our next speaker is Dr. Katie Smith, who will report on the clinical evaluation and she will be followed up by Dr. Catanzaro.

DR. SMITH: Thank you. I am Katie Smith, director of clinical affairs at Gen-Probe. I would like

to give you an overview of the clinical trial results to support this expanded claim for the MTD test.

There has been emphasis on the key analysis and data that we feel supports the expanded claim, and also to address, in a little bit of detail, some of the issues that FDA has raised about the submission of the clinical data and its results.

To begin with, I would just like to give a little background and set the stage for the design of the trial and the data.

First of all, tuberculosis does remain a public health problem, despite the fact that at least in the United States the incidence has declined in the last four years.

A rapid diagnosis of infectious patients is key to rule-in and rule-out TB, so that they can be isolated or not, and treated or not.

In the past, AFB smear and culture and, since 1995 when MTD has become available, these have been the three key laboratory tests to help assess and evaluate a suspicious patient for tuberculosis.

However, these methods have limitations. The specificity of smear is not as good as we would like to see it, because it will pick up every mycobacteria other

than tuberculosis.

The sensitivity of culture may vary from lab to lab. Again, neither test is perfect.

Also, there is an issue when the goal is rapid diagnosis of TB and the time to result. The ~~serum~~-smear result is easily available within a day. However, culture may not be available by any culture method in less than two weeks, if not up to three or four weeks.

Additional background that I would like to provide you with is from a couple of literature references that have come out in the last couple of years.

A presentation made by the CDC at the American Thoracic International Conference in 1998 surveyed TB cases between 1993 and 1996.

Of those 93,437, 13 percent, which is a significant portion, were found to be culture negative. They were TB positive, but missed by culture, in other words.

A more recent article in The Lancet, which I think was provided to the panel members by FDA, studied a specific TB population in the San Francisco area, and determined that 17 percent of those 1,500 cases were smear negative, although culture positive, were, most

important, determined clearly to be infectious.

So, to proceed to the MTD clinical trials to support this expanded claim for smear negative specimens in patients, this was a multi-center, prospective real-time trial at seven individual sites, which were geographically diverse, and also the proportion of patients suspicious for TB varied widely from site to site.

The study design was unique, in that it was based upon an improved approach for establishing a clinical physician diagnosis of TB as an end point.

Prior studies with MTD have looked at the performance versus other laboratory methods such as culture and smear.

This was an improved approach, since it represented a compilation of data and information available to the physician as they evaluate each individual patient.

It also represents the real world of patient evaluation for tuberculosis to date.

The study objectives were primarily to characterize the performance of MTD using the clinical parameters of sensitivity, specificity, positive and negative predictive value.

We also wanted to bring the same measure of performance to smear in culture, using as an end point, again, physician's final diagnosis of TB as the best standard available.

The patient population was important and unique. It represented subjects across all sites who presented with clinical suspicion of TB and who were not on therapy.

Clinical suspicion was based on radiographic findings, laboratory test results such as culture and smear, and also a variety of clinical findings, including signs and symptoms, and a variety of risk factors -- immune status, whether the individual was foreign born, and a variety of other risk factors.

These helped assure there would be a broad spectrum of risk ranging from low to high in each site, and throughout the whole patient study population.

This flow chart represents the process by which each individual patient at each site went, as they were enrolled into the study.

Initially a patient presented at the institution. They underwent an initial assessment by the enrolling or current physician.

At this time, a chest X-ray and tuberculin skin

test were taken, and determination was made as to whether that individual should be put into isolation or therapy.

If that patient met the enrollment criteria, inclusion and exclusion for the study, they were enrolled, again based on a suspicion for TB, as defined by the parameters I just listed. Also, they were not on therapy at that time.

Additional information was gathered, clinical, radiographic and demographic information, and a clinical suspicion of TB infection, based on a percentage ranging from zero to 100 percent, was established by the physician.

Respiratory specimens were collected for smear culture in MTD. Because mycobacterium tuberculosis sheds into the respiratory tract in an unpredictable manner, more than one specimen was collected from each patient who was enrolled in the study.

That does represent current medical and clinical practice. So, each individual in the study had one or up to six specimens that were collected during the course of time they participated in the trial.

Within one to seven days, the MTD and smear results became available, and then within two weeks, or at patient discharge, culture results became available

for the first time, and additional follow-up clinical and radiographic data was obtained, and an updated physician suspicion was established for that patient, that was expressed as a percentage by the physician.

At three-months time, each patient was further evaluated and final culture results were available, follow-up clinical and radiographic data was obtained. Again, an updated suspicion of TB was established.

At the end of this period, a final patient diagnosis was determined by the enrolling physician.

In order to ensure that the diagnosis at each site and for each patient was not subject to a non-uniform basis for diagnosis, because no written established criteria were provided or imposed upon the sites, or the physicians who evaluated these patients, we established an expert panel process to standardize the diagnosis of TB.

The expert panel were independent of the clinical trial, and they established criteria to either rule in or rule out TB.

Each case that was part of the clinical trial was evaluated by this expert panel.

Those that had a clear definition of probable TB, or those that had one of not probable TB, were not

further evaluated by the panel.

However, those that were ambiguous or not clear, according to these criteria established a priori, were further evaluated by the panel.

As a result of this process, 299 cases did not warrant further review panel review by our experts. Forty cases, however, did.

This is significant, because it ensures that at each site, for each patient, the establishment of a diagnosis or suspicion of TB was uniform across all the sites.

Very few patients or cases required further evaluation by the panel.

Based on all the patients enrolled in the study, we determined that 339 were evaluable, and they contributed 834 specimens.

Of these specimens, they were compiled of, or consisted of, both smear-positive and smear-negative specimens. The majority were smear-negative specimens.

This smear by site -- one through seven or noted by APG -- are the number of patients and corresponding specimens contributed to the whole entire data set.

One of the issues that FDA is deliberating on is the ~~portability~~poolability. Initially, all the specimens

collected in a study, and according to the clinical protocol, were tested fresh. That was our intent.

However, it was determined after the study was completed, as we were compiling and analyzing the data, that about a third or a little over a third of evaluable specimens required retesting.

The results obtained on those initially were not scientifically acceptable or valid. So, we wanted to recover those specimens and the data from them.

We did so by presenting a substudy, to establish the answer to one question and that is, are the results on MTD test on fresh specimens comparable or equivalent to those specimens which were processed from isolates, frozen, rethawed and tested. Is fresh equal to frozen.

If the answer is yes, that would allow us to pool results in the MTD test from fresh specimens with those detected on frozen lysates.

We conducted a series of analyses, which I will take a few minutes to demonstrate to you. We determined that frozen was equivalent to fresh.

As a result, we retested specimens that had been stored frozen, and determined that they were combinable or poolable with fresh. Therefore, the data set of 339 patients, 834 specimens, we felt were poolable, and those

represent the patient assessment population to support this claim.

Now, to go into a little bit more detail -- and as I said, fresh and frozen specimen results were pooled -- many of the results were based on an analysis by Fisher's exact test.

Initially, we looked at the data to determine what was appropriate for an estimate of sensitivity, since that was a clinical parameter.

Using Fisher's exact test of P values determined of .261, correspondingly, to address assay specificity, we determined a P value of fresh versus frozen data of .491.

Neither of these P values is significant, which suggests there is not a significant difference between fresh results and frozen results.

Again, this supports our efforts to pool the data and to do analysis on the overall data set.

Furthermore, we wanted to look at subsets of the data. In this case, we looked at only smear-positive specimens.

We determined the sensitivity and specificity in the fresh data set and the frozen data set, performed a Fisher's exact test, and determined, again, that the p

value for either of these estimates is not significant. Therefore, fresh is equivalent to frozen.

The corollary of this, then, of course, is the smear-negative specimen sides. Again, looking at sensitivity in fresh, and specificity in fresh versus frozen, the P values are not significant. The data is poolable, look at the subset of smear-positive and smear-negative specimens.

All of our analyses so far have been based on comparing MTD data with patient diagnosis. However, there was interest on the part of FDA to look at these analyses to estimate poolability using MTD versus culture.

Although this is not our end point or comparator of choice, we did also elect to look at the data in a similar way.

In this case, each MTD result, the first specimen result from each patient, was evaluate versus the first culture result. So, there were paired analyses between MTD and culture on the first specimen from each patient.

The right side shows the results showing fresh specimens, 26 specimens, 26 patients. The left side is the frozen samples.

In this case, when we the analysis was done, the

P values, again, were not significant. They were much greater than .05. Therefore, again, this shows the culture as a measure of MTD in these two subsets of results. The data is not different and, therefore, is poolable.

You can break this out further looking at smear-smear-positive versus smear-negative in the frozen data set, smear-positive versus smear-negative in the fresh data set.

The P values are all either not significant or they are much greater than .05. Again, this extensive further analysis shows that the data is poolable.

DR. CHARACHE: I am sorry, just to clarify this, the cultures are the same? Is it a repeat question, or are these different samples?

MS. JONAS: This represents the first specimen from each patient in the study, and the culture result, the MTD results compared to the specimen, the first one in the sequence of specimens that were collected from each patient.

DR. CHARACHE: So, the culture criteria from the left side and the right side are exactly the same. They are the same cultures, but they are compared to different runs of the MTD test? They have been frozen and run

while they were still fresh.

MS. JONAS: Yes.

DR. CHARACHE: Thank you.

MS. JONAS: So, the question that FDA put to all of us is, can the pool of MTD from all sites be considered relevant to understanding whether MTD is clinically reliable.

We clearly feel that the answer to that is yes.

The data has been shown to be poolable by multiple methods of analysis, comparing MTD with patient diagnosis, and also the culture results.

Standard statistical approaches have been taken, Fisher's exact test and the McNemar's test. It is shown both by patient analysis and specimen analysis.

Again, the package insert, indication for use or intended use, is not directly intended for frozen assays, but we do have a provision in there that frozen assays could be tested if the laboratory chooses to do so.

Now, to get into some of what we feel is the pivotal data to support this expanded claim. This slide is the first one that is summarized in our package insert, for showing the overall clinical trial results.

This represents, again, our evaluable patients out of 339 patients across all study sites. This shows

excellence performance characteristics for MTD compared to patient diagnosis.

Sensitivity is 86 percent, specificity 97.8, a very good positive predictive value and negative predictive value. The confidence intervals for each are shown in parenthesis.

It is always of interest, further, to look at positive predictive value, negative predictive value, as a function of prevalence.

The top line is what negative predictive value there is with prevalence. You can see that it hardly changes at all, when used at a very, very high level.

As you look at positive predictive value, you can see that at lower prevalences it dips a bit, but not measurably when it is determined for all the patients across all the sites. It is actually fairly consistent across a wide range of prevalence.

So, the following chart shows the performance characteristics -- again, sensitivity, specificity, PPV and NPV -- for the entire data set of 339 patients.

Here, one of our other goals was not only to look at and describe that in this study for MTD, but also to see how it compared to other standard tests, namely, smear and culture, and what we defined in our study as

composite culture.

In this case, if any one of the two or three culture methods used by each site in the study was positive, then that particular patient was determined to be positive.

You can see, as you look down the column of sensitivity, that the sensitivity of MTD is vastly better than that of smear, very comparable to culture, or even very comparable to composite culture.

Specificity remains very high for all measures.

PPV is done except for smear, which is done in the other three methods, and is shown here. Again, the negative predictive value is essentially the same.

In conclusion, MTD is very comparable to culture in evaluating these measures. Smear is not quite as good, particularly in sensitivity and positive predictive value.

So, in our clinical trials of 339 patients, we determined that there are 65 who are smear-positive and the vast majority of this group was smear-negative, which was 274.

So, given this, how does the performance of the test, MTD, compare when you are looking at these same characteristics, but separately in the smear-positive

population of 65 compared to smear-negative at 274.

Again, the sensitivity, specificity, PPV and NPV are excellent in smear-positives as would be expected. The smear-negative, the sensitivity is a little bit lower, high specificity, good positive predictive value.

Another evaluation in support is to look at, again, a smear-negative population, the 274 patients. How does that MTD sensitivity, which is lower than smear-positive, how does it compare with the other standard laboratory test.

Well, it is very comparable to the mycobacterial culture methods, very close to, if not comparable, even though it is lower, as I said, in smear-negatives compared to smear-positives.

Again, the specificity is comparable across all these methods. The positive predictive value is quite good and the negative predictive value is very consistent with MTD relative to other culture methods.

The other analysis, as Vivian Jonas mentioned, was to look at an evaluation of the performance of the same measures of the NPV in the MTD tests versus patient diagnosis, we looked at only the first two samples of three or more.

This was done on, not the total of 339, but

approximately 260 of the patients, the criteria being that they all had to contribute at least three or more specimens in order to do this analysis in a balanced way.

As you can see, there is good sensitivity for this whole composition with the first specimen. It is optimized with the first two samples, increasing at three or four specimens.

Specificity is very good across the first sampling. Positive predictive value is very good and negative predictive value is very good.

The other analysis we wanted to see was performed by logistic regression analysis of up to 18 different parameters which reflected on every patient in the trial, across all seven sites.

This was to determine which factor of 18, either considered individually or in combination with MTD, was the most predictive of tuberculosis.

These factors were then ordered when an odds ratio was determined from low to high, and a P value was also calculated to determine if there was any difference.

If an odds ratio for any one of these factors was 1.000, it would indicate that that result provided no more information than not having that result.

If the odds ratio was 500, in the case of MTD

higher than 13, that suggested that it was predictive of prognostic values, when the diagnosis of NPV was 513 times, by having that information or not having that.

So, by far, this says that MTD has tremendous predictive value for tuberculosis and ranks with the clinical and diagnostic parameters.

So, in conclusion, the data that we have shown you, all patients combined, or broken down by smear positive or smear negative, shows that the performance, as reflected by typical and important clinical characteristics, such as sensitivity and specificity, is excellent.

Although the smear-negative sensitivity is slightly lower than smear-positive, we have shown in our study, as mentioned by Mr. Frieberg in his introduction, that of those patients who had TB and were smear negative, MTD picked up 18 out of 27, or 67 percent of this population, which is a significant contribution in the assessment and diagnosis of TB in this study population.

This is another way to depict how each of these three methods -- MTD, culture and smear -- pick up TB in that group of 71 patients shown in the study to have TB.

Again, the 18 that I noted is expressed here.

They were picked up by MTD and culture but not at all by smear.

Forty-two were picked up by all three methods. There is one that is unique in being detected by MTD and smear and one by smear and culture. There are five that only culture picked out, and four in that total set of 71 were not detected by any of the three methods.

Other points to make that are important is that MTD performance does exceed smear and is comparable to culture.

All of the analyses that we have shown you today, and in others that were submitted clearly show this over and over again.

This is just at the base smear or culture, for that matter, but the data clearly show that the population consisted of both smear-positive and smear-negative patients and specimens, the large majority of which were smear negative.

Another consideration is the contribution of MTD relative to the smear and culture for obtaining a rapid diagnosis, since that is an important factor in evaluating a population of patients, especially for TB.

Within our study, MTD results were obtained in a day or less. As I mentioned in my introduction, that is

what is commonly expected for smear.

However, culture, because of the difficulty of growing mycobacterium tuberculosis, takes at least two weeks and often three or four.

So, there is a contribution in the speed to result as well as the accuracy.

Overall, the performance of MTD shows that it is safe and effective for smear-negative patients and specimens. Of those shown in the study to have TB -- which was 71 -- 61 out of 71 that were TB positive were shown and detected by MTD.

Of the remaining 268, 262 were shown to be negative by MTD. This clearly shows that this is safe and effective for smear-negative patients and their associated specimens.

MTD is plainly robust. We have evaluated the data by comparing MTD with a new standard patient recognition diagnosis of TB. We have also look at it by a more traditional mycobacterial standard of culture. The performance characteristics remain the same.

In addition, we have looked at it by comparing it on a patient basis, considering the results of all specimens obtained from all patients.

We have looked at it on a per-specimen basis,

and we have looked at each of those subsets based on patient diagnosis and culture. Again, the performance continues to come out the same, as assessed by clinical sensitivity, specificity, positive and negative predictive values.

Again, while we have shown features in MTD that suggest it could better in terms of smear, we are not intending it to be a replacement.

Smear clearly is important to continue to measure, to estimate the degree of infectiousness. Culture is also intended to be continued and always will be, because it is important to look at mycobacterial susceptibility, which only culture will provide.

MTD, then, is an important adjunctive test in the rapid diagnosis of TB. We feel that the data from this clinical trial does clearly support the expanded claim for smear-negative as well as smear-positive patients.

MR. FRIEBERG: The final presentation is by Dr. Catanzaro.

DR. CATANZARO: Gen-Probe has asked me to come today to make a few comments in response to some specific questions that were raised either by FDA or by themselves.

To begin with, tuberculosis diagnosis is obviously a complex process. Clinicians typically bring together all the information that is available -- history, physical, imaging of various sorts, laboratory and public health information.

Not all of this information is available at the same time. It is an iterative process. You continue to review the data that is available, that has been provided, and make a reassessment.

I think we saw that in the process that Katie Smith has provided, continuing to update the suspicion of tuberculosis.

Obviously, particular for this panel, there is a tremendous emphasis on the laboratory. Clinicians and laboratorians are well aware of both the strengths and the limitations of diagnostic tests, the smear and the culture, and have learned to utilize that information and incorporate it along with the rest of the diagnostic process.

As you well know, it takes many organisms for them to be visualized on the smear, even with the P testing. It is only about 55 percent positive, and the specificity range is quite different from institution to institution.

In some centers, five to ten percent of the smears detect the tuberculosis. In other centers, it can go up to 50 percent.

Culture is frequently felt by many -- particularly laboratorians -- to be the gold standard, and yet, it requires 100 organisms per ml in one colony to be seen.

CDC, in their listing of verified cases of tuberculosis repeatedly finds that perhaps 85 or 90 percent of identified patients are culture positive, so the sensitivity is near 100 percent.

Recently, there has been quite a bit of information suggesting that occasionally false positives occur.

So, all the tests have some degree of sensitivity, some degree of specificity. We have heard, thus far, a fairly extensive presentation of the sensitivity and specificity of the MTD.

I think it is appropriate to put it in context, because as I said, it is a complex diagnosis.

So, the clinical exam has maybe 50 to 60 percent sensitivity, the specificity varies quite substantially, particularly with the experience of the clinician.

Chest X-rays and various studies range from 50

to 75 percent sensitivity, specificity is 60 to 80 percent.

The AFP smear, which we have a lot of interest in, is shown in the next slide. The aspects of the performance of the AFB smear in this particular were the same as for the culture, in this particular study.

So, when we put it in context, then, we have to consider all of these things in establishing a clinical diagnosis of tuberculosis.

There are two points in particular that I would like to address. That is, looking at the data in two different ways.

When patients do not have a diagnosis of tuberculosis, the false positives for MTD, there were, in fact, six out of 268 cases.

I think it is important to recognize that in none of those false positives, could the false positive be returned when the same specimen was tested again, suggesting that there could be laboratory errors.

I think it is worthwhile to consider, what are the potential adverse or down side effects of these false positives.

Patients would be unnecessarily -- there would be unnecessary contact evaluations performed potentially

in these six cases, and potentially, these six cases could be unnecessarily exposed to anti-tuberculosis medications.

So, this is the kind of harm that might occur with those six false positives.

Looking at the other side of the coin, as it were, again starting with the set of patients who did not have tuberculosis, 21 of those individuals had a false positive smear.

The potential benefits of MTD in these patients would have been to avoid potential unnecessary contact evaluations, and avoid unnecessary exposure to anti-tuberculosis medications, so, six versus 21.

The patients who had tuberculosis, on the other hand, looking at false negatives first, with the MTD, this occurred in 10 of the 71 cases in this study, who had a final analysis and diagnosis of tuberculosis.

Five of those 10 patients were, in fact, negative on culture. They were among the class of culture-negative tuberculosis.

Two of the patients had samples that had inhibitors, and one of these 10 was actually positive by AFB smear.

What are the potential adverse or downside

occurrences from these false negatives? Well, we have potential transmission of tuberculosis to these cases. It could be a source of tuberculosis transmission.

I would point out that the number is 10, and that all but one of those was smear-negative cases.

On the other side, if you look at patients who had tuberculosis, the true positives with MTD, 61 of the 71 patients had TB. Twenty-seven of those were negative on the AFB smear.

So, the potential benefit of using MTD in these cases was to avoid the potential transmission of tuberculosis in these 27 cases that were smear-negative cases, and the potential benefit of starting effective anti-tuberculosis therapy earlier, interrupting transmission and starting patients on their way to care.

The other potential benefit of the MTD is one of the ruling out tuberculosis. Patients obviously come to the hospital because they are sick, and not because they are worried that they may have TB. Those patients obviously have something wrong with them.

This is the distribution of the other diagnoses that were established in the cases. The majority, this number is 72, 71 had tuberculosis. The rest had infections other than tuberculosis.

I want to draw your attention to the 27 who had, in fact, pulmonary neoplasm. I think MTD offers some clear advantages to these individuals. They can have the correct diagnosis established a bit more promptly perhaps if MTD is used to make a rapid diagnosis or a rapid ruling out factor in these patients.

Finally, I would like to give a little bit of attention to the problems in defining the clinical suspicion of tuberculosis.

This is a rather difficult task. I know that there is some focus on this, both in this trial and in the use of the rapid diagnostic test.

CDC has done quite a bit of work in this area, identifying epidemiologic risk factors, to focus clinicians' attention on the likelihood of a particular individual being at risk for having active tuberculosis.

This isn't the same at all as establishing the clinical definition and likelihood of tuberculosis. The likelihood of tuberculosis requires the assimilation of the points that I have already mentioned.

There was a workshop on examining the issue of what is the appropriate use of rapid diagnostic tests for tuberculosis.

At that workshop held in San Diego, 120 experts

in TB were brought together to consider this point.

They felt that it was important to focus the rapid diagnostic test on patients who were suspected of having tuberculosis, but were unable to come up with any specific definition of how to define the risk of TB and the suspicion of TB, which then had a high likelihood of tuberculosis.

We have a situation that is kind of awkward. Clinicians clearly know what a TB suspect is, they clearly know what a highly suspicious case of tuberculosis is, but no one seems to be able to define that very well.

In fact, when we brought together a group of clinical investigators to undertake the clinical trial that was presented a few minutes ago, that happened after the workshop.

One of the important points that we tried to do was to bring the clinicians together, to focus in on patients who were suspected of having tuberculosis.

As I said, we were unable to come up with a definition. In fact, CDC is currently conducting a utility trial and is having similar problems.

So, doctors know the diagnosis of tuberculosis, and this test will be adjunct to the other measures that

have been identified -- history and physical, X-ray, laboratory, public health -- all of that comes together with the MTD acting as an adjunct that can be used at various steps along the way. Thank you for your attention.

MR. FRIEBERG: To conclude, the purpose of our application is to add the five words to the intended use.

Addressing Tony's last point, we have a current approved use for patients suspected of having tuberculosis. We do not request that any change be made to that.

We would like to reiterate that this is adjunctive to other tests. The sensitivity and specificity of the product has been demonstrated by patient diagnosis. We believe that is the best method of proof in this particular case.

We have addressed labeling warnings for the expanded use, along with data in the proposed package inserts, which you have all received.

Therefore, we believe we have addressed the FDA issues to modify this particular intended use, adding smear-negative data to the package insert, so that we have used sufficient, least burdensome clinical evaluation methods to bring this before you. Thank you.

DR. CHARACHE: Questions for the sponsor?

DR. O'BRIEN: Maybe I should introduce myself, since I wasn't here this morning. I am Rick O'Brien. I am head of the research and evaluation branch in the TB division of CDC in Atlanta.

In the spirit of full disclosure, I should mention that staff in our branch have been involved with Gen-Probe in several studies, one completed and one just being undertaken and one still in the discussion stage, for which Gen-Probe will be providing and has provided the test kits.

I have not been involved myself, directly, with the company with these discussions or in planning the studies. These facts were disclosed to FDA before I came on the panel.

Mr. Frieberg, you mentioned that 18 patients who were smear negative and MTD positive could have been more efficiently diagnosed and treated more properly, had the test been even used.

I assume you have information -- maybe you have it available -- but how many of those 18 patients actually did have presumptive treatment begun prior to the culture results being available?

Frequently, patients with negative smears have a

sufficiently high clinical index of suspicion to have presumptive treatment begun.

That might be of interest in determining whether or not your suggestion is correct.

DR. CATANZARO: I don't have the specific results, but I want to point out that the MTD results were never available to the clinicians.

So, all clinical judgements were made a priori, without that information.

Further, as was pointed out by one of the speakers already, there is a great tendency -- in fact, in this study, there was a great tendency for the MTD to be considerably more accurate than the institutional of chemotherapy for active tuberculosis.

While I don't have an accurate answer in numbers, maybe we can pull that out.

MR. FRIEBERG: It was not available.

DR. CATANZARO: It is not available, it was not a factor.

DR. O'BRIEN: My question was -- and the data are in your data set, or at least they are on the commission evaluation forms. How many of those 18 patients had presumptive treatment initiated, even though they were smear negative.

How quick were the clinicians in these sites at beginning treatment for smear-negative patients, which is done?

MR. FRIEBERG: Gail Woods is suggesting she should address that.

DR. WOODS: My name is Gail Woods. I am at the University of Texas Medical Center in Galveston. I honestly cannot address your particular question with regard to this particular data set.

I can tell you that, in an article, in a study that we performed after the clinical trials were completed, and the data had been published, we looked at I believe it was 1,004 specimens from close to 500 patients. I don't remember the exact numbers.

Anyway, there were 22 patients in that group, all of whom I guess I should say were Texas State offenders. We wanted to focus on that particular population of patients because they are at particularly high risk of TB, being in an enclosed environment, might benefit from the results, although they did not get the results.

The bottom line is, of the 22 patients, 10 of them were smear positive and 12 of them were smear negative. Several of those smear-negative patients were

outpatients.

In the inpatient group, I looked at the charts and, in two of those patients, it would have made a very significant difference had that result been available.

They were HIV-infected patients. One of the patients -- both were smear negative. The first specimen from each patient was MTD positive.

One of the patients was started on therapy a couple of days after the sputum specimen had been collected. Regardless of being started on therapy, they were not convinced enough that the patient had TB to just leave it at that. They had to do an invasive procedure.

In the other patient, they did not anti-tubercular therapy. So, he went for I believe it was about two weeks without therapy.

In one patient, it was a minimum of \$12,000 hospital costs, not including physicians, and the other patient, a minimum of \$22,000.

At least in our particular situation, I think that those -- they are isolated cases, but I think they can justify, at least for our population, that in those types of situations the test would have been very valuable.

DR. EDELSTEIN: Dr. Woods, may I ask a question,

while we are bringing up costs? I seem to recall in your study that there were several people who had false positive actions as well.

DR. WOODS: This is true.

DR. EDELSTEIN: What was the cost associated with that?

DR. EDELSTEIN: As I think I said, but let me make it perfectly clear, the physician did not get the MTD results. So, therefore, there was no cost associated.

I want to address these false positives.

MR. FRIEBERG: We really shouldn't be addressing costs.

DR. WOODS: I know, and I apologize.

DR. CHARACHE: Let's hold this for just a moment. I think the issue of the false positives is a very real consideration. How we address that, if we have it come forward through the overall panel discussion, it is labeled here as a cost, but the whole issue of false positives, I think, will have to be thought about and we may have our discussions following the sponsors based on that.

MR. FRIEBERG: We want to make sure that the issue is fully addressed, but it sounded like he was

asking a practice of medicine question. Is it useful? That is not a safety and efficacy question.

DR. CHARACHE: This question, where the laboratory's responsibility begins and ends, and where the clinician's responsibility begins and ends, I think we do have to ask a couple of questions about it, because that has been an emphasis of multiple people from the sponsor's group.

Here, I think you had a very specific question in terms of the results of a false positive. Can we pick that up as we talk about false positives, and then we can talk about the positive predictive value in your institution versus others? Would that be all right?

DR. EDELSTEIN: That is fine.

DR. CHARACHE: Any other questions from other people now, of this panel?

DR. O'BRIEN: I don't know if it is a question or a comment.

DR. CHARACHE: No, comments will be later. I would like to ask for a further clarification of the separation that has been drawn between the clinician's responsibility to establish a diagnosis and the role of a microbiology device in assisting in that decision making.

I think here, I would specifically ask if I am

understanding the concept being presented, that the laboratory does or does not have the responsibility to select a test, or perform a test, based on whether it is appropriate for the population at that institution.

Are we saying that only the clinician can make that decision, or what are we saying?

MR. FRIEBERG: I don't think we are changing anything at all. Whether or not our TB test is run now, with smear positive or smear negative, that decision making process is in place. We are not changing it. By adding smear negative, we are not changing the practice.

DR. CHARACHE: I think there is a perspective that the interpretation of a laboratory test, that any laboratory that performs a test should be prepared to explain what that test means.

If they can't assist the clinician in understanding the meaning of the test, they shouldn't be performing it. Do we have any disagreement on this?

Then another related question is that, if the prevalence of the given disease doesn't warrant the testing in that laboratory, is it the responsibility of the laboratorian to say that this test should not be performed, or does the clinician have the responsibility to say, I want it done anyway.

I am not sure where you are saying the practice of medicine and the practice of laboratory science, in all targeted patients, merge or separate or exactly what you are saying here.

MR. FRIEBERG: I believe that is the regulatory side. We try to regulate it by statements in the package insert, but I don't see that that would be respected in the field. That is not something that is regulated.

DR. CHARACHE: That pertains to safety and efficacy.

DR. CATANZARO: I am a clinician, obviously, and maybe that is the problem here, having the clinician stick his nose into the laboratory.

I think there is a great deal to be gained by clinicians and laboratories working together. I think that all populations, patients are important as well.

It is certainly possible, on a clinical basis, to define, from a whole set of patients, the small set that has a high incidence of tuberculosis and the larger set that has lower incidence.

I think the laboratorians and the clinicians should work together, is number one.

Number two is that CDC is famous for making many guidelines that help clinicians and laboratorians work

together.

While the comments are made that FDA doesn't have the responsibility, clearly, CDC does, and has exercised it regularly, in conjunction with the American Thoracic Society. I think that is an appropriate place for these kinds of judgements to be made.

MR. FRIEBERG: We can hear from Dr. Woods on the subject as well.

DR. WOODS: I would like to second what Dr. Catanzaro just said. I truly believe that the laboratories and physicians should work together. That is the way it has always been with my laboratory.

I agree with what you said, that if the incidence prevalence in a particular place -- and we could pick on any particular state, and there are several where the prevalence is quite low -- perhaps the laboratory director decides that it is not cost effective to offer the test in that particular institution.

Well, maybe there is a particular patient in whom tuberculosis is suspected and a clinician feels very strongly that he or she would like that test.

I think if it were me and the clinician were coming to me to say this, I would feel obligated to send that out to have it done.

I have been told this by many clinicians, when they are mad at me because we don't do what they want, I am not there taking care of that patient. It is not going to be my responsibility if they die, and blah, blah, blah.

I do feel very obligated -- that is a true statement. I am not there taking care of that patient. Therefore, I respect their opinion. If they feel very strongly that this is what needs to be done, unless it is a totally absurd thing, which occasionally happens, and I can't talk them out of it, that is what we would do.

DR. GUTMAN: I agree with Dr. Frieberg, when he suggests that our purview does not involve the practice of laboratory medicine or the practice of clinical medicine.

I perhaps disagree with him on the notion that the labeling issues are not relevant to making sure that the product is effective and safe, so that it can be used in whatever appropriate laboratory or clinical practice.

We certainly -- as you will see from the questions the FDA proposed -- are concerned with how to label a product like this, to make sure that people practicing in various ways, both laboratorians and clinicians, hopefully working together but not always

working together, will be able to understand and use the device.

DR. CHARACHE: I also was not making a break between those who carry a stethoscope and those who don't. I think you will find a number of laboratory directors here, including me, who also carry a stethoscope. We are not making that break.

I was really thinking in terms of safety and efficacy and how that gets translated into saying that the person who orders the test knows how to use it, because we know that that is not necessarily the case.

DR. SMITH: I just wanted to make one other comment to address your question, Dr. Charache, and that is the study design was such that at each site there were co-investigators or co-study directors, one coming from the clinical side and one coming from the laboratory side.

So, there was the collaboration in the study and conduct of our trial just like Dr. Catanzaro and Woods were describing their experiences.

DR. CHARACHE: While you are up, I have a question for you, and then I will try to elicit other questions.

You showed the sensitivity of the frozen versus

the fresh panel. I did just a little bit of quick calculations.

It looked to me as though the positive predictive value was not the same. I wonder if you could tell us about the positive predictive value about the smear negatives.

DR. SMITH: I think there is a little less performance overall in that particular parameter, relative to that. It is a little bit lower in the frozen set versus fresh.

DR. CHARACHE: This is what I calculated; that there was not a lot of positive predictive values.

DR. SMITH: So, that would suggest that there are more false positives in the frozen set. The best that we can ascribe that to would be the additional handling of the specimens, in terms of processing them after they are stored, frozen, thawed and retested, that there may have been some contamination that was introduced.

DR. CHARACHE: I think this may be of interest, because the whole issue of contamination is very important if you are going to merge the frozen and the fresh, when it comes to the culture negatives.

I am focusing on the culture negatives. Much of

the data, the majority of the data was positives and negatives. I am trying to pull out of that, that which pertains to the smear negatives. That is what we have to address.

It would be very interesting to know what the positive predictive value changes were in those.

DR. SMITH: Again, it wasn't the vast majority that did undergo that change or that handling dilemma or problems.

DR. CHARACHE: Maybe we could just get a look at it. That is one of our questions, which is, can they be pooled. When predictive value changes, that raises a question.

MR. CHEN: I am Dafeng Chen, a statistician from Gen-Probe. You said that positive predictive value -- we have done many, many analysis. We have used the first specimen, the first two specimens. We used culture as the reference test and patient diagnosis. I don't know which numbers you used.

DR. CHARACHE: The one that was just presented. That was the culture. It was the same culture with the test run twice, so you had the solid data base to compare it to.

MR. CHEN: For smear negative.

DR. CHARACHE: Yes, that is what I am asking. Other questions for the sponsor?

DR. EDELSTEIN: I have several questions about the poolability of the data, and the fresh/frozen, and trying to validate this approach and analyzing the specimens that have been frozen for, I think, about a year before they were retested.

You performed a study in which you compared test sensitivity and specificity on a different population, using samples that had been stored for seven, ten, up to 30 days before re-analysis.

In the comparison, between the fresh analysis and the 30-day analysis, the specificity of the test decreased from 96 percent to 88 percent.

What I would like -- and I have some questions, number one, about the significance of that change, and I think it might be that the appropriate statistic to use for that set is the McNemar~~EMR~~ analysis, which I haven't seen, because I haven't seen that. I would like to know which ones changed.

The second question is, how were the specimens stored at all of the study sites for that one-year period? Were they stored in phosphate freezers? Were they stored under constant monitoring? Was there

documentation that there were no thaws that occurred during shipment, all those sorts of questions.

MS. JONAS: In answer to your question, how the specimens were stored, they were stored at -70 at Gen-Probe, so they were not stored at the individual sites.

The testing was completed, we got shipments of them in house. We stored them at Gen-Probe, and all them are under constant monitoring alarm systems and everything, so no, there was no freeze/thaw.

DR. EDELSTEIN: How were they stored at the study sites before they were sent to Gen-Probe?

MS. JONAS: They were aliquotted into -- they remained in the lysing tube and they were stored at -70 the whole time.

MR. FRIEBERG: -20 to -70.

DR. EDELSTEIN: For those that were stored at -20, were those freezers monitored?

MS. JONAS: At the sites, yes. It was a short time.

DR. EDELSTEIN: Can you provide me with the [EMR McNemar](#) analysis of that data set, please?

MR. CHEN: We have conducted analysis to show that there was no difference between the fresh and frozen. The P value is slightly below .05 for fresh

versus frozen.

DR. CHARACHE: Any other questions?

DR. WEINSTEIN: I want to go back to prevalence of disease. At our hospital, between 1993 and 1998, the prevalence of TB was 3.6 percent.

At the East Orange VA Medical Center, in a more urban environment, the prevalence during that same period was 3.1 percent.

When I look at your graph, page V-7, book 3, my concern is what happens to the positive predictive value when you are in an institution with prevalence that is well below five percent.

It looks to me, based on the graph, that one out of every three positives would be a false positive.

DR. CATANZARO: I think you know that we are not suggesting that we use it at any institution, rather that we use the test for patients who are suspected of having tuberculosis.

At our medical center, if you simply monitor the incidence of positive cultures that are sent to the laboratory, it is about eight percent. It is a little bit higher than yours, but it is still rather low, certainly compared to incidence of tuberculosis in patients who were enrolled in studies.

This is where clinical judgement comes into play. Certainly, as the prevalence goes up, the positive predictive value improves. That is where physicians play a role. If you test a whole city, you are going to get a lot of false positives. We certainly don't want that.

You are absolutely right. No matter what curve you work with, as you have this decrease, the value changes.

MR. FRIEBERG: We are not suggesting that we should publish prevalence everywhere.

DR. WEINSTEIN: My concern is that, if you use the test as you are proposing to use the test, without providing any guidelines in your package insert, some information about prevalence, there are going to be institutions that may use this test in a way that will cause one out of every three patients, in our example, to be treated erroneously.

DR. CATANZARO: That was the purpose of my presentation, to discuss the difficulty in coming up with specific guidelines for how to diagnose tuberculosis.

CDC, ~~NTS~~-ATS and others have worked quite hard and have not been able to do that. We don't think we can do that in the package insert.

DR. CHARACHE: I think we should proceed with

the FDA, and ask the sponsors to remain, so we can ask questions later, too.

**AGENDA ITEM: FDA Presentation.**

DR. SIMONE: Hi, my name is Patty Simone. I have been asked to briefly review the impact of NAA testing on TB treatment and control programs.

Briefly, the role of the smear in TB control is that we use the smear to influence our clinical suspicion of tuberculosis. We use it to help decide on the probability of infectiousness, and we also use it to indicate the response to the patient's therapy.

Culture is used -- a positive culture nearly always confirms a diagnosis of tuberculosis, but you can have a false positive culture, just like you can have false positive other things.

In different laboratories, the range is from one percent up to maybe eight percent, for an average of four percent false positive cultures.

A negative culture does not preclude the diagnosis of tuberculosis, as we have heard earlier.

Positive culture allows susceptibility testing.

Also, it may show the growth of other mycobacteria which will give an alternative diagnosis, and cultures are also used to follow response to therapy.

In our current MTB control, a TB suspect is reported to the health department. The ratio of suspects to cases can vary quite a bit across the United States, anywhere from two to one up to six or even more to one, in different parts of the United States.

based on the report, a contact investigation is initiated, nearly always for smear-positive patients, and some health departments do not begin initiating this with patients who are smear negative. Further diagnostic testing is performed and then treatment is started.

The TB case is confirmed either by a positive culture or, if there were TB signs and symptoms present that improve with treatment. So, either of those two things will usually go into the diagnosis of TB.

The third practice in infection control is that TB isolation is initiated based on a variety of things -- symptoms, smear, clinical presentation.

The criteria for an initiation of isolation varies greatly, by the prevalence of TB and TB risk factors in a community, the type of facility, isolation capacity, et cetera.

In some parts of the country, there may be as many as 20 to 1 patients isolated than actually had tuberculosis and in others it is much lower.

TB isolation is discontinued either when TB is ruled out -- usually when we have three negative smears and another diagnosis or explanation for the clinical syndrome. In this case, we do it after a minimum of three days.

If the patient is actually confirmed to be a TB case, isolation is discontinued when the patient shows adequate clinical response, including negative smears and clinical response to treatment.

This may take up to two weeks. However, patients are often discharged home earlier when they become seres smear negative at the hospital.

When we have our initial results of our evaluation, a smear positive result will trigger a contact investigation, which will trigger the start of treatment, and isolation is either initiated if it hasn't been done, or continue to be initiated, based on a clinical diagnosis.

A smear-negative result initially may instigate a contact investigation, although many health departments wait until the culture has been done, in smear-negative cases.

Treatment may be started if clinical suspicion is high, and isolation is usually not continued if it was

started initially, when the patient becomes smear-negative.

Once we have the culture results, a smear-positive, culture-positive patient, the contact investigation is completed by the health department and we need to ensure that the patient completes the full course of treatment.

If the patient is smear-negative and culture positive, then if we haven't already done a contact investigation, we may do so, and we need to ensure that the patient completes the full course of treatment.

If the patient is found to be smear negative and culture negative, the contact investigation usually is not done, because the likelihood of infectious disease is really low at this time.

The treatment is continued if there is clinical improvement. It may be stopped if there is no clinical improvement, or it may be modified to deal with a particular infection rather than TB.

What would be the impact of a 100 percent incidence of infectious disease. If you have a patient who is smear positive and NAA positive, these are the basically the patients who are potentially infectious with TB, we basically have no impact, because we were

already going to treat this patient, put him in isolation, and start a contact investigation.

If we have a patient who is smear positive and NAA negative -- for example, patients who have non-tubercular kinds of bacterial disease -- an NAA test may prevent or reduce unnecessary contact investigations, unnecessary patient treatment or unnecessary TB isolation.

I know we are not supposed to talk about cost, but these are all very costly and resource-intensive activities, and basically this covers that.

If a patient is smear negative but NAA positive -- that is, they have TB but maybe it is less infectious -- an NAA test may help reduce delays in the initiation of therapy for patients who were smear-negative, when there was low clinical suspicion. If there was high clinical suspicion, that therapy would have been started anyway.

If there is a patient who is smear negative and NAA negative -- that is, if their culture was also negative -- if we had a high suspicion of TB, we may consider a therapeutic trial. If we had low suspicion of TB, then we would most likely to consider an alternative diagnosis.

What are the implications of a false positive result in TB prevention or control. If there is a false positive NAA test in a patient who is smear negative, it may make the clinician start treatment they may not have started otherwise. It may instigate some contact investigations that may not have been done otherwise.

Again, false positive NAA test, and a smear positive patients, then some unnecessary contact investigation may be done, although many would already be should started anyway, some unnecessary isolation would have been done. This is very similar to clinical practice that is currently done.

In a smear-negative patient, if you have a false negative NAA result, there would basically be no impact if there was low clinical suspicion.

However, if there was high clinical suspicion, it might influence you to delay therapy.

If the patient is smear positive and has a false negative NAA result, this may delay treatment, contact investigation and isolation that the clinician may have started having not had false negative results.

I just wanted to briefly review some of the relevant issues in TB control. The first is that TB already is decreasing at a very creditable rate of at

least five or six percent per year. This, of course, will affect the positive predictive value of a test, when the incidence goes down.

Also, it reduces clinical expertise, and makes the interpretation and complex decision making much more difficult.

It also reduces laboratory expertise and proficiency, if there are fewer specimens being processed.

The final thought from a public health standpoint is that we have level or reduced federal funding for TB prevention and control, and we always must be looking for things that are cost effective -- not just helpful, but cost effective, because we have less funding to take care of the problem.

MS. SHIVELY: I would like to thank Dr. Simone for coming to talk to us this morning. I am the next presenter for the FDA's part this afternoon.

Before we get started, I would like to thank the other members of the review team for this PMA supplement application.

Also I would like to thank the sponsor for the presentations and the information that was presented, the clinical study design and the application.

As Mr. Frieberg pointed out, this is a PMA supplement application for a new intended use fof a device. As such, FDA considers this to be a significant~~submitted~~ change to the product that is already commercially available.

This afternoon, we are going to focus on the study design, the data, the applicant's data analyse~~is~~, along with statistical considerations for those analyse~~is~~.

We consider the evidence from the sponsor to be important for determining the effectiveness of the MTD for this new proposed intended use.

Also, we feel that it is important to identify patient populations for use for the MTD and to indicate and, if necessary, contraindicate for harm.

Also, the evidence provided by the applicant is necessary to have adequate understanding of the product's performance with its new intended use, to be able to provide adequate directions for use, including precautions, warnings, interpretations and guidelines for interpretations.

As background, the sponsor has already described some of the marketing history of the MTD, and it is currently approved for use with AFBP smear-positive

specimens.

I have included a summary scheme of the interpretation guidelines that are contained in the current package insert for the MTD product used for AFB smear-positive specimens.

As a note, warnings and precautions regarding the performance of the MTD in this population are incorporated into these interpretive guidelines.

Thus, for an MTD's negative result for a smear-positive specimen, it should be considered to have either MOTT, either MOTT plus MT**b** or Mtb could be present but inhibited.

Also, additionally, there is a recommendation to test another specimen, if the patient is suspected to have clinically active TB, or if inhibition is suspected.

The applicant's proposed new use does include a few words indicating a use for smear-negative specimens.

FDA believes that this wording implies use of the MTD for any specimen from an untreated patient with a suspicion of TB, or for TB in a differential diagnosis.

We also believe that the wording of this proposed intended use suggests that the MTD may be the first laboratory evidence for confirming a clinical suspicion of TB.

Also, we would note that the applicant did use the same interpretation guidelines for AFB smear-negative specimens in patients as for currently approved smear-positive specimens.

FDA believes the new intended use should be based on the applicant's evidence and the patient population evaluated in the applicant's study.

We would recommend alternate wording for the intended use. First, we would consider presumptive as a qualification for the level of laboratory evidence provided by an MTD test in a smear-negative patient.

Also, we would want to qualify a description of the target patient population to something that brings into consideration the clinical suspicion level of the patients being considered for testing.

We would like to thank Dr. ~~Kapazara(?)~~ Catanazaro, that maybe likelihood is a word to incorporate.

Additionally, there have been various discussions about the influence of prevalence. FDA certainly would be interested in the perspective that the panel has on the influence of prevalence for use of the MTD.

Because FDA believes it is essential to

interpret MTD in conjunction with AFB smear findings, we would include an additional statement in the intended use, that a concurrent AFB smear is necessary to interpret MTD results on selected patients.

The next area we will look at is specific interpretive criteria for this new target population for the smear-negative testings.

FDA does agree with the applicant that the smear-positive component of the insert should remain unchanged, and certainly, laboratories may choose to continue to use smear-positive as a selection criteria for performing MTD testing.

In this slide, we have shown possible modifications of what could be considered for the modified interpretation guidelines, for use with smear-negative specimens.

For an MTD-positive result in a smear-negative specimen, we may want to consider that MT**b**Ⓟ has the lower likelihood of actually being in the specimen, although it could be probable or possible.

Certainly, MOTT could still be possible, and there may be value to an additional test from that patient, to verify the initial MTD-positive result.

For an MTD-negative result, MT**b**Ⓟ may certainly

be unlikely, but it may also be present in low numbers or inhibited, and there may be value, again, for doing a second specimen on a patient with the smear-negative specimen, that tests MTD negative.

I am moving on to number nine. The sponsor has provided a very good overview of the study design of the study that was performed.

I would like to note that the patients that were evaluated in this study, study eligibility was based on a low to high suspicion of TB.

Criteria are different by site and by practice of clinician, or perhaps it would be better to say that they were not specific criteria.

We do not know if there are different criteria used at the different sites. We do know that overall, from the 339 patients, 80.5 percent were placed in isolation, 29.8 percent were begun on multi-drug therapy, 74.6 percent had cough of minimum duration, 95.9 percent had abnormal chest X-ray findings, 30.4 percent were HIV positive, 19.2 percent were smear positive.

The applicant describes the evaluative population by physician suspicion. The FDA would prefer more objective criteria for defining these patients.

Again, we are certainly interested in this

panel's perspective on how this patient population ~~responds~~can be defined, or whether suspicion is an adequate indication.

We did note, that the ~~Harbor-UCLA(?)~~Francis J. Curry National TB ~~Center~~study does provide a guide to the related criteria for triage, and that this may provide useful information.

To understand the applicant's study population relative to the smear status, this chart shows the ~~extent to which~~percentage of the smear-positive and smear-negative populations ~~go~~ across the study sites, and also across overall study populations.

The smear positive portion of the study population was 65/339 patients. ~~These is~~ are is the red lines at the top of the column with the number designated at the top. Smear negative are the blue.

Looking across sites, that is the total number ~~off~~ patients at each site.

At the site on the far left, of the 13 patients that were evaluated, three were smear negative. At another site, here in the middle, of the 50 patients who were evaluated, 48 were smear negative.

Prevalence is always an important consideration for a study population. For all the patients evaluated,

which is the red column, 21 percent were TB positive.

Of all the smear positive patients that were studied, 66 percent were smear-positive. Of the smear-negative, 10 percent were TB positive.

We note that this is the percentage TB positive, using the applicant's definition of physician/clinician diagnosis.

One of FDA's review issues has been to characterize the applicant's patient study population relative to proposed target populations in the intended use.

The applicant describes the study population basically based on a definition of clinician suspicion criteria, and FDA has discussed whether this is an adequate criteria for this type of population and, more importantly, would conclusions from~~prompt~~ an evaluation of such patients could be applied to a selected population or applied to a broad spectrum of patients in any clinical setting and laboratory facilities.

The next area we were looking at are FDA considered issues of the applicant creat~~developing~~ a relevant data base that would provide appropriate evidence to support the package insert intended use modifications.

First, we would note that the number of MTD tests per patients varies. Seventy-seven patients had four or more MTD tests, 260 patients had at least two MTD tests, and 179 patients had only one test.

Two hundred ninety eight, out of 884 MTD evaluable results were from retested frozen lysates.

Thirdly, FDA would be interested in using the applicant's recommended interpretation criteria that is incorporated ~~ed~~ ing the second specimen MTD results.

Fourth, we looked at the TB/No TB patient status categorization. We do think that this would be an adequate characterization to use for the intended use. However, we would note that there is some variation in that definition that, from the expert panel review, if considered, some ~~off~~ the categorizations would ~~for~~ change.

Also, the question was posed as to whether pulmonary should be differentiated from extra-pulmonary TB, particularly in the culture-negative patients. Lastly, we are concerned with potential ~~side~~ side effects.

Given the types of data from the applicant's study, FDA does believe that the first specimen analysis ~~offers the package~~ avoids the bias effects of multiple samples, and also serves to maximize the data available, by including all 339 patients.

However, unfortunately, it doesn't represent recommendations for additional testing, particularly for those patients who would be considered to have clinically suspicious TB.

Our first plus second analysis is an approach that we have looked at, and this would use the second MTD test result, if the first MTD test was negative.

We believe that this would be consistent with the conditions for use and the directions for use of the MTD. It assumes that all of the first MTD negative patients would have clinical suspicion.

I got a little lost in my wording here on the slide. I want to say that we want to specify that where the 2nd MTD test is available, we would consider the patient to be clinically suspicious.

However, when a second MTD test was not available, there is an uncertainty as to how to assign that patient.

In considering whether MTD results from tested frozen lysates should be used in a primary analysis, FDA does believe that the MTD was intended to be used during the initial evaluation of the patient, and that the benefit of the MTD test is for physician waiting for diagnostic evidence from culture, particularly for smear-

negative patients.

I will finish up here and turn the presentation over to John Dawson, our statistician. However, before we do that, I would like to point out to you all that a data base that has been assessed that FDA has developed, based on sponsor's evidence, is included in your package.

In this data set, we are taking the first MTD result if it was evaluatable~~available~~, and used the smear status information on the patient against~~fter~~ the culture status of the patient, and defining culture-negative TB patients separately.

Do you have any questions on this data base? Using it, we proceeded to do several different types of analysis, and then John, would you like to follow through on these at this point, or would you like me to go ahead?

MR. DAWSON: Why don't you go ahead.

MS. SHIVELY: In this first analysis, which is the first MTD analysis, sensitivity of the MTD for the smear-negative population was 59.4 percent, specificity, 98.8 percent.

We did note that we consider culture negative TB patients to be in the TB positive category here. We have also shown the smear-positive analysis, but.—~~Because we believe that the proposed new intended use~~amendment~~ is~~

specifically for smear-negatives, we are focusing on that population.

We did include all 290 of the patients who had an initial smear negative result.

A second analysis was also done using a second MTD result incorporated with the first, the rule being that if the first MTD result were negative, that if a second MTD test was available, we would use that.

This again, used the entire 290 smear-negative patient grouping. Sensitivity was estimated to be 68.8 percent, but it improves over the first analysis. Specificity was 97.7 percent, a decrease.

I put little arrows here to indicate where the movement was between cells, using this statistical analysis.

One consideration with the first MTD approach, was what to do with those patients who didn't have a second specimen, or a second MTD test available.

We did a variation of the previous analysis. We dropped those patients who had no second specimen ~~available~~evaluable for the smear-negative population. The sensitivity using this approach was 75.9 percent, specificity 97 percent.

I now will turn this presentation over to John

Dawson, FDA statistician, to discuss some of the more relevant statistical concerns.

MR. DAWSON: Good afternoon. I am John Dawson. I am the statistical reviewer on this application. I am the last speaker.

I heard recently, at a conference at CDC, that going last is bad, except in ice skating and Russian Roulette.

I want to go further with the definition of a working data set, that Roxanne was talking about, and I want to talk about some analysis of that data set.

Looking for a subset is a little bit unusual, particularly for an FDA statistician who criticizes companies for doing this.

There were two reasons for doing it in this situation. One is that the intended use, as you can see, is still under negotiation, and has been since September.

The other reason is that itthere is probably not intended use to find out, ~~in the last particular,~~ that both the company and the FDA are going to agree on that will useusing all 834 specimens and all 339 patients.

Let me just quickly go down the list of the things that we were looking for in the way of a working data set.

First, we wanted to be able to use all the specimens. I put this first, because the data that I am going to show you doesn't have fresh only, it is fresh plus frozen.

We felt this was important, because this could be the normal intended ~~use, even though the indication is for frozen specimens.~~

On the second point, we did want to be able to look at smear positive and smear negative separately. Thirdly, we wanted to eliminate any cases that had reagent problems.

Fourthly, we wanted to be able to use a second MTD test when the first one was negative.

Next, we wanted to be able to consider all cultures, not simply the first culture, and regard any culture positive patient as being diseasepatient positive.

We also decided to omit, after some considerable discussion among ourselves, 67 patients that were negative on the first MTD test, and did not have a second specimen.

This gives us a subset of 269 out of 339 patients, and utilizes 485 of the 834 specimens.

As to the analyses, we wanted to use MTD to

evaluate -- rather, we wanted to be able to assess~~use~~ its ability to predict clinical status.

By clinical status, I mean in terms of a definition of operational proof, any culture positive or culture negative TB patients, of which there were, I believe, five that had a diagnosis of TB.

We wanted to do the analysis on a per patient basis, rather than per specimen. We wanted to use the cut off of 30,000 relative light units, and to use the retest result of equivocal cases.

Again, this uses both fresh and frozen specimens. This is basically a quality control slide, from my point of view.

~~Th~~This deals with the 45 patients in our working data base who were smear positive, and I will explain what is in this.

These are the results by site and results overall. You have just seen this from Roxanne. The numbers in the lower right-hand corner are, ~~for~~ the totals across all sites.

To read this, for example, for Brooklyn, we have three cases that are positive by clinical status, and MTD also. There is one case that was negative by both.

I say quality control in the sense that, we have

reason to believe that this test works well in smear positive cases.

I wanted to look at this to be sure that there is evidence of MTD working well in smear positive cases.

I hoped to find not very many discrepant and, as it turned out, there were zero discrepant.

If this has shown poor performance in smear-positive cases, I would have been very worried about our working data set.

This slide is~~is~~ the heart of it, as far as I am concerned. There are~~is~~ the smear-negative cases, fresh and frozen specimens, 224 patients, 129 specimens.

We no longer had the situation where there were no discrepant specimens. For example, in Galveston, there are two cases that are positive by clinical standards status and MTD also.

There are 23 cases that are negative~~later~~ by both, and we had three discrepant~~discovery~~ cases.

I wanted to see three things. One is, can we pool across sites. I wanted to see what the ability of the MTD in this working data set was, in terms of what whether we could agree with and predict clinical status.

I wanted to obtain performance estimates for MTD

with relative to clinical status.

First, as far as poolability, I looked at it from the point of chi square tests of homogeneity, of sensitivity. ~~A~~ across sites. ~~It~~ It is essentially a two-by-seven table. It is ~~software that is~~ equivalent to kind of a big Fisher's exact test.

It passed that test. It also passed the test looking at specificities. I also looked at the McNemar test, for the reason that the McNemar provides a sort of conservative way of evaluating the agreement between two methods of making a diagnosis. It doesn't depend on a gold standard. It simply compares the rates of positive response.

As for the comparison, the table that we have just looked at -- and again, these are the smear-negative cases, and looking only at the total across all sites table, 95.5% of cases are on the main diagonal, which I considered to be a pretty good result. You see the confidence interval there, and Kappa correlation is .74.

In McNemar's test of equal rates of positive response, it has a P value of 1.0, which is about as high as you can go.

For the comparison to clinical status, this has the ability to predict status. Status determination is

two weeks to a month, or six weeks or three months off into the future. ~~Is it able to pass the threshold of time for a given patient.~~

We have already seen that the McNemar was not significant, which is favorable, indicating comparable rates for both MTD2 and ~~for considering~~ clinical status, as if it was another method.

I just wanted to back this up with the odds ratio, because the McNemar only considers the discrepant cases. As one of my colleagues pointed out to me, you can back up McNemar with the odds ratio ~~as using which~~ uses results from all four cells from the table.

It had an odds ratio of greater than 1.0, which indicates that. ~~T~~there is a predictive value. So, ~~looking at~~ the odds ratio and the McNemar together indicates the performance of the MTD2 test relative to clinical status in that working data set.

Performance estimates for this data set, the sensitivity 76 percent, lower confidence ~~limit interval~~, 57 percent, specificity 97 percent, lower confidence limit, 94 percent.

Finally, I just want to indicate my concern about the problem of having the fresh and frozen in the same data set. This side of the table shows the frozen

specimen cases, this shows the fresh specimen cases on the right.

I draw your attention to what happens with the smear negative for the fresh specimens. We are getting down to a very small subset of cases, but there is an imbalance.

This was based on a data set that does not consider the second MTD, in situations where the first MTD is negative.

It is possible, and we speculate that, in some cases, the second MTD may resolve the discrepancy, what may appear to be a false negative based on the prior MTD, but I don't know that. That is just speculation~~ing~~.

In the frozen subset, the smear negative, there is greater balance. In terms of McNemar, this is not significant, which is good. ~~It~~This is significant in the fresh subset, which is bad.

I have three conclusions. One is, the results suggest the ability of MTD to predict clinical status.

Secondly, it is unknown what it would be with the fresh specimens. If, in fact, using fresh and frozen combined in the same data set makes sense from a clinical point of view, then we have that accounted for.

I feel if fresh and frozen are both the intended

use, then I think ~~there should be~~ these numbers provide a reasonable set of performance estimates.

Thirdly, this is after ~~the~~ fact subsetting, and warrants confirmation of results, though not necessarily ~~in the~~ post-approval confirmation. I would be happy to see it in pre-approval.

Failing ~~any~~ most of those, I think the labeling should indicate that the results have not been validated.

Thank you very much.

Dr. Charache: Thank you. Do you have questions to present?

MS. SHIVELY: Yes. FDA would also like to go through the questions that we have prepared for the panel discussion today.

First, the applicant proposes to pool MTD data from retested frozen lysates, which were done by a separate laboratory, with MTD from testing fresh specimens at six out of the seven clinical laboratories.

Can the pooled data be used to characterize performance for individual sites?

If yes, how should this data be represented on the labeling.

B, is the data -- fresh, frozen or pooled --

adequate to characterize individual site performance for the use of this device.

C, if not, what types of data or primary analysis should be used for laboratory site performance evaluation?

The next question. The table referenced here has been included in the copy which you have in your packages.

The question states, should the instructions produced in the label include information to clarify differences in expected performance for smear-negative versus smear-positive specimens.

If so, where and how is this information best meaningful.

B, if not, are any other guidance or caveat in the label appropriate to ensure safe and effective use of the MTD for smear-negative specimens.

The last question, the first part of the question is: Does the current study, plus data and information from previous studies, provide sufficient evidence to modify current labeling as requested of the applicant.

If yes, does the panel have recommendations for other labeling modifications, such as the

contraindications, warnings or limitations, to ensure safe and effective use, or request a change to those guidances.

As a subset of this question, should MTD testing of smear-negative specimens be indicated for selected patients, such as those with high clinical suspicion.

A second part of this option, should labeling explicitly link use of the MTD to test smear-negative specimens in high prevalence settings.

The second major part of this question, if no, what additional data or data analysis might be appropriate to support the requested intended use modification.

Are there other alternatives, such as labeling modifications, that could support MTD use for smear-negative patients.

DR. CHARACHE: Thank you. Are there questions for the FDA speakers before we proceed?

Hearing none, this morning, the sponsor was provided a five-minute comment period after the FDA presentation, before the open public hearing.

I am questioning whether the sponsor would like to make a response at this time, or whether you would like to proceed.

MR. FRIEBERG: We will respond.

**AGENDA ITEM: Industry Response.**

DR. SMITH: I just wanted to revisit -- I showed this data earlier. It is not new. It is comparable to an analysis that Mr. Dawson presented, but it has a difference also that I wanted to emphasize.

All the results here are fresh versus frozen, smear-positive versus smear-negative. We have shown the P value to not be significant. Therefore, they are poolable. They are the same, and they don't have to be subsetted.

One of the things that we did here that was different from his analysis is that, my understanding is that the best application of the McNemar test is where you have paired results. That is why I introduce this.

Each MTD result was compared to a culture of the same specimen. The specimen was always the first specimen for each patient.

The comparison of MTD to culture in all these subsets is controlled, although it was only in the first specimen.

I believe his analysis included the results of all the culture specimens on that patient, compared to only the first MTD result. So, there was a little bit

more scatter for culture compared to MTD.

Whether that results in a difference in the P values, which I don't know, but we feel this is a very rigorous approach to looking at the data and using the McNemar test, which is a conservative one for looking at this data in this way.

DR. CHARACHE: just one point, while that slide is up. That was the basis for my questioning the predictive values. The predictive value from the frozen one would be 57 percent, and the predictive value from the fresh would be 91 percent.

That is why I was wondering why that occurred, what that meant.

DR. SMITH: Well, I guess two factors. One I already mentioned, that some of the specimens, we felt, did switch values after being thawed and retested. We contend that this can happen with the fresh specimens.

The frequency of that is very low. The only thing is you have a difference in the total number of specimens in the two sets. Again, I am not a statistician. So, that is all I can say.

We feel this is a valid analysis, and also justifies the pooling of the data.

MR. FRIEBERG: The only reason that we wanted to

put it up is because of the statement that the tests are not poolable. We have data to show that they are comparable.

MR. CHEN: Another factor, what drives the difference in the predictive value is the prevalence. If you look at the prevalence for testing the fresh specimens, you have the prevalence, if you add all the culture positive together, you have 14, versus the culture negative 125.

So, the prevalence is 14 over 125. In the fresh data set you have 42 plus five. That is 47 over 177. So, it is the difference in prevalence which also drives the difference in the predictive values.

MR. FRIEBERG: One more comment from Dr. Catanzaro, and then we will close, to try to keep it under five minutes.

DR. CATANZARO: Dr. Simone did a really nice job of presenting the analysis from the standpoint of TB. I want to remind you of the comment that I made, and that is that most of the patients in this study came to the hospital because they were sick, not because they thought they had TB.

Of the 338, 71 turned out to have TB. The vast majority, 267, did not have tuberculosis. I think the

best use of this test is, in fact, in the smear-negative cases.

I think the best use of this test is to rule out the diagnosis of tuberculosis. I think that is further substantiated by the fact that, of the six false positives, none of them were repeated. They could be confirmed by a second test.

Furthermore, in this trial, 38 percent of patients were put on therapy. Eighty percent of the patients in this trial were put in isolation.

This explains why TB costs us a lot of money to take care of. In fact, the majority of the funds that are to take care of TB -- the majority of the funds used to take care of TB -- are in fact, spent on patients who do not have tuberculosis.

In the study done by UCSD, we demonstrated that 67 percent of the dollars that we spend on TB are spent on patients who do not have TB.

I think that any test that allows us to quickly move patients through that isolation -- as Dr. Simone pointed out, in some centers, 20 patients are identified as TB suspects for every one who turns out to have tuberculosis.

The quicker we can move those folks on to the

direct diagnosis, we have done a tremendous service to those patients.

I think the calculations that are done here, to add up both columns and divide it by the number of patients that are diagnosed is missing the boat. The value is in finding patients who do not have tuberculosis, in general and in the study.

DR. WOODS: I would just like to address one of the suggestions that was made with regard to -- I apologize for not being able to exactly reproduce what you had on one of your slides, it was with the ones -- maybe you could just refresh my memory.

When you had a smear-negative, MTD positive, you suggested that one of the things that should be stated is that a second specimen should be considered, and if there is only one, the results should be considered inconclusive. Was it something along those lines?

MS. SHIVELY: Yes.

DR. WOODS: Just based on the data from the one study that we did, although granted, it is not included in the clinical trial, we did have, as we told you before, 10 patients who were smear negative and both culture positive and MTD positive.

Of those 10 patients, three had only one

specimen. For the five of the seven patients, for whom there was more than one specimen, only one specimen was MTD positive.

Therefore, by adding those additional requirements, those patients could have been considered to not have tuberculosis. This could happen.

Therefore, my suggestion would be to not make such a dogmatic statement. It is not a bad idea to have a number of specimens, but if you then discount the results --

DR. CHARACHE: I am going to have to interrupt. I was going to wait until you finished and then say that we have to strike that, because that is included in the kind of data that we are not allowed to review. It hasn't been presented earlier.

DR. WOODS: So, this is only clinical trial data.

DR. O'BRIEN: They handed out the paper.

DR. CHARACHE: That was in the handout?

DR. O'BRIEN: It is in the handout.

DR. CHARACHE: That is wonderful, but the five minutes is over. No, keep going. I am just alerting you. Just continue your thought.

DR. WOODS: The question of false positives was

brought up earlier and you alluded to the fact that I might be able to re-address it and I was wondering whether this might be one of those opportunities. But if not --

MR. FRIEBERG: Let's just conclude by saying that anything can be applied well, anything can be applied wrongly.

Our approach to applying the test well is as the intended use will show you, and with sensitivity and specificity.

We believe we have shown that the sensitivity and specificity is different for smear negative, and is quantifiable for the package insert.

I am very concerned about quantifying the prevalence across the city, across the country, and putting it in the package insert.

I am also concerned about certain definitions, that suspicious is objective. Highly suspicious is also subjective.

We found that we could not quantify highly suspicious. That is why we believe we should leave the intended use as it is, leave it up to the clinical team to define what suspicious use is, and that it is appropriate to use the test given the data we provided.

DR. CHARACHE: Thank you very much. Let's have a brief break, return at 4:10, and we will continue with public comments.

[Brief recess.]

DR. CHARACHE: I would like to continue. Before continuing with any public discussion, we would also like to permit the FDA discussants to provide any additional information or clarification, if they wish. We also had cut them off earlier. In the interests of fairness, could we hear from the FDA?

**AGENDA ITEM: FDA Response.**

MS. SHIVELY: I would just like to make one clarification. Dr. Woods made a comment at the end of the last session.

Regarding the interpretation of guidelines that could be considered for possible modifications for the indicated smear-negative population, FDA is suggesting that these are options that could be considered, but that they should be based on evidence provided in the applicant's study.

We certainly wouldn't consider repeating a second specimen unless we had the data to support doing that, or not to do it, too. Thank you.

DR. CHARACHE: Thank you. Any other comments or

additions? Thank you very much.

There have been no individuals who have, in advance, requested to make comments during this public comment period.

We would ask if there is anyone here from the public who would like to make a comment at this time.

Hearing none and seeing none, we will continue with the open committee discussion. We will not have another break before it.

**AGENDA ITEM: Committee Discussion,  
Recommendations and Vote.**

DR. CHARACHE: We would like to begin the committee discussion with the individual who did the primary review of this for the panel, and that is Michael Wilson. Will you lead us?

DR. WILSON: Thank you very much. The request for the change in the labeling is to extend the indications for this test to patients with negative smears.

This raises several questions about the data set that I had, and I would like to go through these very briefly, and then open it up for discussion.

One that I am somewhat concerned about the validity is patient diagnosis as the gold standard,

against which one is comparing the performance characteristics of a laboratory test.

This is using a somewhat different standard than has been used in the past for other tests, particularly in this case where there were no defined criteria for patient diagnosis, and for using a gold standard test in which there are no defined criteria.

I think that it is unlikely that such a criteria would pass muster, for example, for publication in a peer-reviewed journal. Essentially, the criteria used in the gold standard remain undefined.

The study concern I had is that what we are really concerned about is the care of patients who have tuberculosis, pulmonary tuberculosis.

We do want to exclude tuberculosis in patients who do not have it, but in those patients who do have it, the total number of patients that we are talking about here is 27 patients.

The N for this clinical trial is really 27 patients. I think that is an insufficient number on which to base changing the labeling indication.

The third problem that I think we have to address has been raised before, this issue of prevalence.

While the clinical suspicion in a given patient

is important, the prevalence is based on the patient population from which that patient is derived.

As we have already seen in this before, as the prevalence drops, the positive predictive value of this test also varies significantly.

The last concern that I had is that at several of the clinical sites there were very small numbers of patients.

I have concern whenever we have that few patients, and also the proficiency of the testing done on that site.

As we have already seen, at least in one publication, there is an attachment to Dr. Woods, that this is a very trainer-dependent test.

The person who does not perform this test in adequate numbers is not going to have the training and is unlikely to perform it well. So, I have some concerns about the proficiency and individual test competency on this.

Lastly, I do have concerned about the sensitivity, even though I understand why the sponsor is more concerned with specificity.

The truth is that the sensitivity of this test on a smear-negative patient with pulmonary tuberculosis

is only a little over 50 percent, which is not much better than a coin toss.

So, I would like to make those initial comments, and then participate with the rest of the formal discussion.

DR. CHARACHE: I hear you raising questions about the validity of the study, as a basis of making the decisions that we have been asked to address.

I would like the view of the panel. Should we consider the questions in the order in which they were presented, or should we first consider this underlying issue?

I think all six of the points we have heard about pertain to whether this study provides the valid basis required to decide on package wording changes.

I think I would like to ask the preferences of the panel and their guidance in terms of how you would like to proceed.

DR. WEINSTEIN: Just review with us again how you came up with the N of 27?

DR. WILSON: On page 51 of the handout, this is smear-negative patients, there were 27 patients who were diagnosed with TB who were also smear-negative patients.

DR. GATES: Unless I got it wrong, listening to

both the FDA and the sponsor statistician talk about the data, they both thought, from what I can tell, that the data was sufficiently correct. They both agreed that the data was accurate.

I guess I am also thinking, from the point of view of the FDA modernization act, in terms of least burdensome data, the idea is that as long as the data are sufficient to say whether the test meets its claims or not, that it should be the least amount of data that is necessary to do that and still maintain good safety and efficacy standards. In my mind, it seems like it is okay.

DR. CHARACHE: I think perhaps if we elect to discuss the efficacy first, this may be one of the first points. How many patients do we base this type of decision on.

DR. O'BRIEN: It sounded like you were calling for a vote on whether we go through the questions in order or address the studies. I don't know whether you want to actually do that, or continue discussing the comments by Dr. Wilson.

DR. CHARACHE: I think Dr. Wilson had this as a framework. Do you wish to amplify this?

DR. O'BRIEN: To take a somewhat contrary

viewpoint, I think the study design, in general, was quite good, and that the point the company made, that using the expert clinical diagnosis as the gold standard for something like this is quite appropriate.

I have been involved in TB studies where we have done that, because the laboratories aren't perfect. There are eight percent false positive culture results in culture negative TB and 10 to 20 percent.

Using clinical diagnosis, which heavily relies on the laboratory, I think, is very appropriate.

Even though it wasn't done, there was some specificity given to the diagnosis. All patients who had clinical symptoms who were suspected of TB and had two positive cultures, were considered tuberculosis patients.

DR. CHARACHE: I think what I am asking is, should we have this discussion first?

DR. O'BRIEN: Oh, I understood you to have said yes.

DR. CHARACHE: No, I want the panel to decide whether we should -- we can go through the points raised by Dr. Wilson one at a time, and decide whether they are important or not, or what we think the importance is, and then go through the questions, or we can go through the questions first. I would like to see what the panel

prefers to do.

DR. SANDERS: I think we should address Dr. Wilson's concerns first and then have questions.

DR. CHARACHE: We have one suggestion that we address the concerns first and then go through the detailed questions. Other points?

DR. EDELSTEIN: I vote just the opposite. If we go through the questions, as we perhaps discuss those questions, we can bring up Dr. Wilson's questions dealing with the data set.

DR. CHARACHE: Okay, so we now have two thoughts on the table. Let's hear some more.

DR. O'BRIEN: I would opt for having a general discussion before addressing the questions.

DR. CHARACHE: Shall we go around?

DR. TUAZON: I think as we go through the questions, I think we would be able to discuss specific points that Dr. Wilson has.

DR. SPECTER: It is hard to go through these questions and address Dr. Wilson's points, too.

DR. RELLER: I think the questions provide the framework to address the issues Dr. Wilson raised.

DR. GATES: I agree.

MR. REYNOLDS: Do the questions.

DR. WEINSTEIN: I am with them.

DR. CHARACHE: We have consensus that we should go through the questions and address the questions raised by Mike as we go.

So, we will come back to your questions later, if it is still an issue later. Can we see the first question, please? If we have overheads, we can get more light.

The first question is, the applicant proposes to pool MTD data from retested frozen lysates done at a single laboratory facility, with MTD data from testing fresh specimens at six out of seven clinical laboratories. Can this pool data be used to characterize performance for individual sites. That is our first question. Any comments or suggestions?

DR. TUAZON: I have a question. Why did we need to use data on frozen lysates, when, if we do the test, we use fresh specimens?

DR. CHARACHE: Let me ask the manufacturer what harm would be done by withdrawing the frozen lysates, because they do add issues.

DR. SMITH: I guess our preference is not to do that for two reasons. One is that it diminishes the data set and the significance and perhaps the results drawn

from that.

We actually have analyzed only the fresh data. The performance of the test is excellent. It is the same as we see in the combined data set, and fresh than frozen, for the most part, we found a little bit better.

We feel it is more valuable to include all the data, because we have shown it poolable, and it would give more significance to the user of the test if there is a larger data set.

That portion of the package insert that would break out the data by sites would be more representative than if we only used the fresh data. Our preference is to use the pooled data.

DR. TUAZON: Would the clinician ask for the test with the -- would they be done in frozen specimens?

DR. SMITH: No, it is expected it would be done on fresh. Again, if you agree or buy into the fact that we have shown frozen retested specimens give you the same results as fresh specimens, then the need to distinguish fresh or frozen goes away.

It is no longer important to ongoing use of the test. Again, the data set and the performance of the test would be supported to a greater extent by the pooled data set than just looking at fresh.

DR. TUAZON: On the other hand, the applicability of the test is for rapid diagnosis. You would like use a fresh specimen.

DR. CHARACHE: Am I correct, that since there were seven smear-negative specimens that were in the frozen pool, that would reduce the total number of smear-negative, culture-positive specimens to 20?

DR. O'BRIEN: Seventeen.

DR. MATHEWS: The fresh data set, we did not break it out that way. That is not what we were concerned about.

DR. CHARACHE: They were on Dr. Smith's slide. Oh, I see. Well, we can get that from the data. Dr. Edelstein?

DR. EDELSTEIN: I think what it does for the data, essentially, the performance of the fresh specimens alone for the smear-negative patients actually looks a little bit better. It widens the confidence interval at the estimate of the performance.

I haven't done those calculations, but my guess is, with only 17 as opposed to 26 or 27 patients, it broadens it quite a bit.

I think that is one thing we have to look at, is to look at not only what the estimate is, but what the

confidence interval of that estimate is.

DR. SMITH: Thank you, that illustrates my point, that by reducing the sample set, in order to just express and demonstrate the fresh data, you will probably broaden the confidence interval.

To the user, going forward in time, they will not have as much confidence that that is representative of the test in general, or either in their population in their site.

Fortifying the data set with a larger number will help ensure that confidence intervals will be narrower, and that the performance of the test will be more accurately described now and in the future.

DR. CHARACHE: Thank you. I had a question on this point. That pertains to that particular hypothesis, that perhaps you could have contamination of some of those samples, giving rise to the lower predictive value and the higher culture negative.

MS. SMITH: I would like to just show you some additional data.

DR. CHARACHE: We can't.

MS. SMITH: We can show the impact of the data set, if we just look at the fresh. There is the initial data set, which was obtained with good and not-good

reagent, and compare that with the pooled data, and fresh and frozen. The performance characteristics don't change.

Although there was an incidence of seven going from negative to false positive, the impact of that, switching quadrants in the two-by-two table, did not really diminish the test performance over all. I appreciate your concern about that.

DR. CHARACHE: Okay. Dr. Reller?

DR. RELLER: At the end of the day, we will be interested in how this test for the purposes of the supplement, how this test performs on smear-negative samples from patients who have tuberculosis.

Given the guidelines, the questions that Dr. Wilson had would be tackled in this framework. What I would like to ask is, inclusion of the frozen sample results doesn't increase the number of patients with TB who had smear-negative results; right?

I mean, it doesn't augment the data base for making an ultimate evaluation of the performance of this test in smear-negative patients.

What was the purpose -- not whether the results are altered one way or the other -- but what was the fundamental reason for freezing the samples in one

central place and retesting them. What was the purpose?

MS. JONAS: We have to admit something a little embarrassing. During the course of the clinical trial, we discovered that one of the reagents wasn't purified properly.

Once we found that out, we stopped the trial, called back all those reagents. Especially in those sites who hadn't completed their study yet, they had to stop until we got the new reagents.

We have been through this with the FDA. They are aware of it. We have instituted new QC procedures and this isn't going to happen again.

The problem is that there was an amount of data -- and I don't have the numbers off the top of my head -- that has already been generated with the bad primer.

We did not want to use that data, because you can appreciate scientifically why you wouldn't want to use impure primer in the study.

Based on that, we had saved all the lysates. First of all, we wanted to do inhibition testing. Second of all, we wanted to have it and reserve it in case something happened. In this case, something did happen.

We decided that, rather than burden each clinical trial site -- and there were reasons why some

clinical trial sites couldn't retest no specimens anyway -- we contracted with yet another site, Massachusetts General Hospital, Mary Jane Ferrar's lab, to do that testing.

It was unfortunate. If anything can go wrong with MTD, it does, trust me, in a clinical trial.

DR. RELLER: My concerns on pursuing this frozen data is the implications for how the efficacy of the test, as it is intended to be used.

It is okay to freeze the specimen, but it depends on how long they were frozen and what are the data we have for that.

Given the way testing is going in this nation, one is making an assessment on efficacy one way, but some of that was derived from frozen samples, and you can freeze them and shift them off to anyplace.

Coming back to Dr. Wilson's central issue, what I would like to know is how many patients with tuberculosis had negative smears and their specimens were tested a test like that which is being marketed now with the same requirements.

If our smear-negative patients with confirmed tuberculosis, is there a requirement of good primers, bad primers on fresh specimens, good primers on frozen

specimens, some of which were stored at -20, some of which at -70, but all sent swiftly -- but we don't know how swiftly swiftly was, or in what proportions and how those match up with patients with TB with smear-negative -- you see what the questions are.

I know where I hope we are ultimately. The question is, are we there now.

Do we have enough data to make those decisions objectively about the efficacy? How many patients are we talking about and how was the testing done? I am not sure any more.

DR. CHARACHE: Dr. Reller, that kind of risk that you gave, is that kind of information required for you to be able to resolve this question?

DR. RELLER: Yes, I think central to the proposed alterations for extension of this test to smear-negative patients, that one must have, for efficacious use, a reasonable understanding of which smear-negative patients this would be effective to use on, and what to do when the test is -- what it means when it is positive on smear-positives, and the current algorithm that is there, and what one needs to do when the test is positive or negative in smear-negative patients.

The prevalence questions that come up are very

important. I totally agree with Gen-Probe that prevalence varies, prevalence changes, and prevalence has a major effect, and is crucial for positive predictive value and negative predictive value.

It does not affect sensitivity and specificity, which is what I am trying to get the emphasis on, in terms of the patients who really have the entity.

What you do have, that is important, is what the pretest probability is or not, and getting some way for the individual clinician, in working in concert with the laboratory, to decide what it means when you get a positive on a smear positive, and what it means when you get a positive or a negative on a smear negative.

To have the proper sensitivity and specificity, I need to know what the performance was with reagents on fresh or frozen specimens, you know, from patients who really had the disease, who we are interested in diagnosing.

DR. CHARACHE: Other thoughts and comments?

DR. GATES: I think one of the fundamental questions is, is the data that we see here predictive of how the test is going to work in the lab.

Then, the question of using frozen or fresh, does that introduce any bias. We see that the data we

have seen indicates that it doesn't.

I mean, this has been pretty rigorous. But fresh specimens and frozen are equivalent in predicting how the test will work in the real world.

From a different point of view, I guess, from the industrial point of view, although I am not involved in any way with the test, it isn't that different from working on any TB-type test.

Believe me, it is not easy to get the clinical samples that you need to do it. It is a tough test to do. I think one of the questions is, is the data that we have now sufficient, given the fact that there are not that many tests on the market and there is clinical efficacy.

That is, there is clinical utility with various tests on the market, given the fact that it looks like the data is equivalent.

Is it worth re-doing the clinical testing to have more data, or is this enough data, given the fact that there are not that many tests in the market, and there is a clinical utility for this already.

DR. CHARACHE: I think our first concentration is, what is the satisfaction of the panel and what is their advice, on whether the frozen sample can be pooled

with the fresh sample, based on information that is now available, or would additional information be required in order to do that.

DR. EDELSTEIN: I can express an opinion. I am not convinced that the data sets can be pooled, but the effect of not pooling them doesn't affect the estimate of the test performance. It only broadens the confidence interval.

What I would like to propose is that the confidence intervals for different patient prevalences be published in the product insert, so that people could determine what the test performance would be for their patient populations.

I think the confidence intervals are broad and some people may not view the test as being that good, but overall, I am not convinced that the data can be pooled, because of the question of what happens to the specimens, with the storage -- we don't really have data.

We have a data set for 30 days. We don't have a data set for a year, to see what happens to test performance.

DR. RELLER: I am specifically concerned about the question of frozen samples. It doesn't seem to me that the study design was intended to provide data for

this question.

I mean, my understanding is that the frozen samples were all done with the same primer, but the original specimens, unfortunately, were not done with the same primer; is that correct?

Consequently -- I mean, you do fresh specimens with different primers. I am sorry, but I think you have two things going on on the one, and one thing going on on the frozen, and put them all together so you have tight confidence intervals. It doesn't make any sense to me.

What about the numbers?

DR. CHARACHE: I think Dr. Reller has said that he is concerned about the data. Dr. Edelstein has said that he wouldn't pool, but he is not quite as concerned about the data.

DR. EDELSTEIN: No, I am not saying that. I am saying I agree with Dr. Reller that the data should not be pooled, but the practical significance doesn't affect the estimate of the test performance. It only broadens the confidence interval. I agree completely that they should not be pooled.

DR. GUTMAN: Can I just make sure I understand this, because what you are saying, I am not sure I am following.

It seems to me that the argument that the company is making was that if you pooled the data, you wouldn't broaden the confidence interval; you would narrow it.

DR. EDELSTEIN: That is right.

DR. GUTMAN: Okay, fine. The other point --

DR. EDELSTEIN: What I am saying is that if you only consider the specimens that were tested without the faulty primer, then the confidence interval would be broadened.

DR. CHARACHE: Would you like to say anything further?

DR. WILSON: No, I agree with Dr. Reller and Dr. Edelstein, that I don't believe the data can be pooled.

DR. CHARACHE: Should we just take a moment to go around and see if anyone would like to volunteer whether they think it can be pooled, should not be pooled, or they don't feel they want to comment on the particular question?

DR. HAMMERSCHLAG: I think I concur. I don't think it can be pooled. Further, I don't think that we can really use it to characterize individual sites.

It is such a heterogeneous population and

populations of such variance in numbers from site to site, and the prevalence of TB and everything, that it is very difficult to extrapolate necessarily the performance at one site.

I think the major problem to begin with is how to define populations of people. I don't think many people know very much about their characteristics, even of their own populations.

Then when you start with these tests being used in a large commercial laboratory where specimens are coming in from all over the place, I don't know how one is going to be able to deal with that in terms of the population characteristics, because they are not going to deal with the whole population.

DR. SANDERS: I agree that the data should be pooled, and it also raised the whole issue of what is the actual N that we are dealing with in these smear-negative patients.

DR. WEINSTEIN: What was the third option?

DR. CHARACHE: To say you would rather not comment.

DR. WEINSTEIN: Okay, no comment.

MR. REYNOLDS: I, again, with the rest of the group, have some reservations about the use of frozen

specimens, with no data showing the stability of these tests, no comparative study.

Really, it is almost like you are comparing apples and oranges. I am not comfortable with using the frozen data.

DR. GATES: I think, for reasons I have said, the data that I have seen from both sides is that you can have a sense of what you are doing, with how it works in the field.

DR. CHARACHE: I think Dr. Reller has spoken. Dr. Specter?

DR. SPECTER: I have no comment.

DR. O'BRIEN: I haven't heard anything that would definitely convince me that they couldn't be pooled. It seems that the data provided by the company suggested that, on the basis of the tests done, that the results should be equivalent.

I was concerned, though, about the numbers of patients that might be available for the analysis if only the fresh specimens were used, and specifically the smear negatives.

It gives you 17. I just wonder how we can come to any conclusion on the basis of 17 patients.

DR. TUAZON: I think they should not be pooled.

I think when we did this test, we would be using fresh specimens, not frozen specimens.

DR. WILSON: I have already said, I don't think they can be pooled.

DR. CHARACHE: So, two of our members think that it is okay, and one did not comment -- I am sorry, two did not wish to comment.

DR. GUTMAN: I guess you can move on now to section C, since you have answered that they can't be pooled. Then the question is what types of data or further analysis -- no, I guess we go to number two. If they can't be pooled, then the question Dr. Wilson asked is, if you don't allow the frozen or pooled to be used, is the data for fresh adequate to characterize the individual site performance.

DR. CHARACHE: Actually, this is point number three raised. Even including the pooled data, there was concern expressed about using 27 individuals, and now we are down to 17. I would like to hear some thoughts.

DR. EDELSTEIN: I think that excluding the re-analyzed specimen, while the data could be used to characterize it, the estimate of what the true value would be, would be very broad.

I think it would also be very broad, even if the

data were pooled, because of the tremendous heterogeneity between the different data collection analysis at sites.

I would say no to fresh, frozen or pooled.

DR. CHARACHE: So, you are saying that the elimination of the frozen group would not modify your -- why don't you say it.

DR. EDELSTEIN: I would say the pooled data is inadequate to characterize the individual site performance, just because of the small numbers at each of the sites for the samples of interest, which are the smear-negative, culture-positive specimens. Even at the largest study site, there are very few data.

DR. CHARACHE: Other thoughts and comments?

DR. O'BRIEN: Just thinking about TB trials and practical considerations, these are the most difficult patients to find for studies.

We may be asking something that is not so feasible or desirable from a cost viewpoint. It may be that the ideal studies are difficult to be done or unlikely to be done to satisfy us.

DR. EDELSTEIN: Then maybe I misunderstand the question, but it is my understanding that this is to characterize the individual site performance, not to use the aggregate data. That is different.

DR. CHARACHE: I think there is a need to show consistency between sites, so that we know that different laboratories will get results that can be analyzed.

Other comments?

DR. SANDERS: This really isn't in regard to this question, but going back to the question that I think was asked under issues.

Gen-Probe already has the indication for smear-positive patients. They are asking for an expanded use for smear-negative patients.

There is nothing out there to preclude a clinician or a laboratory from running this test on a smear-negative patient at this time.

DR. CHARACHE: It would be against the --

DR. SANDERS: It would be an off-label use of the device, essentially, is what it would be, if it is being done.

However, if we are seeing patients who we think have TB, even if they are smear-negative, and we are going to do culture, that is already indicated, that this is an adjunct to diagnosis.

I wonder if we are being more burdensome than we really need to be, to answer the question of the expanded use. I just raise the issue.

DR. CHARACHE: So, you are pointing out that laboratories sometimes do off-label things. What we are being asked to address here is whether the label can be changed.

I think all of us who enjoy that particular organism are very sensitive to the fact that you have to do an incredible number of negatives to get your pool of culture positives in negative patients, that there is a small subset of that, as has been made clear.

I think the question on the table now is, understanding the difficulty of getting the information, and understanding that some practitioners are creative, the question has to be, is the data there on which this decision can be made at this point.

DR. GATES: Just as a point of clarification, the data that was submitted to the FDA before it got to the panel. Is there any indication that the design was flawed or there wasn't enough data at that point? Is this the first time the data been analyzed in terms of its specificity?

DR. GUTMAN: We framed this with all the issues, or all the salient issues identified up front. I want to focus -- what you see is what we got.

The issues we identified were the issues. I

don't think we were overjoyed by the numbers. I don't know that we were horrified.

We are trying to titrate against the least burdensome threshold. That is the law, and we do appreciate the unique difficulty in acquiring these positives.

That is what you are being paid the big bucks to do, is to come here and quality control us, and to ask the questions that we didn't ask and think of the things we didn't think.

So, it is absolutely -- the fact that we didn't identify it doesn't make it a wrong question or a right question.

We have a diverse crew that is looking at this submission, with a variety of different ideas about how the data presented, and tried to summarize it in our presentations.

That wasn't a question. Our first question was, are there enough numbers here. It may be a secondary question, that if you refuse to pool, it makes it a scanty number set. I am not shocked, surprised or horrified that you raised the issue.

You do have to keep in mind, from what I presume you folks were instructed this morning, is this is hard

to collect large numbers of specimens.

What I was getting at was, the point that Dr. Edelstein was making, if you look at the data, if you go back and look at it, if you did pool, you would give the company the benefit of broader confidence intervals, but you would penalize their point estimate, because their performance, at least in the data we have seen, deteriorates in the frozen.

So, you give them better confidence, but you give them lousier performance. We think, looking at the data set, we understand why the performance deteriorated in the frozen.

This may be magical thinking or a rationalization, but we think that the sensitivity fell because the RNA was labile and we think that the specificity fell because contaminations were raised.

I am not going to prejudice you one way or the other. You are supposed to prejudice me.

DR. SPECTER: At this point, I really am not sure what we are talking about at all, in terms of what is the data set. So, I picked my own to look at.

What I picked was, I believe, was shown by John Dawson of the FDA, which is the MTD2 present clinical status, which is as reasonable a data set as any of them

here, based on statements made by some of the people.

If one looks at that particular data set and looks at the smear negative --

DR. CHARACHE: What page are you looking at?

DR. SPECTER: I have no clue what number it is.

DR. CHARACHE: It is at the bottom.

DR. SPECTER: It is actually two pages in front of where our questions are. If you look at the bottom and you choose only fresh specimens, which people wanted to look at, and only smear negatives, you will see that there are 136 different negatives on the MTD2 if you look at that last little two-by-two.

Then you see that there are 11 positive by the MTD2, six that were negative by the MTD2, there were quite a few positives, and there were no false positives according to clinical status. That is a data set that one could look at.

If one looks at that particular data set, then one could try to make some conclusions, perhaps.

DR. CHARACHE: That is fresh or frozen.

DR. SPECTER: No, it is fresh only.

DR. CHARACHE: I am sorry, I am looking at the wrong page. It is actually comparing the data between fresh and frozen.

MR. DAWSON: We can put that chart up.

DR. SPECTER: In looking at this, I see that, if one likes to use clinical status -- and Dr. O'Brien has indicated that is a very useful thing to do -- and this test says it was useful in 11 additional patients and there were no false positives, this would suggest that there is benefit here.

The question is, then, is this a reasonable data set to use, or is it not, because we are talking about so many different options, I don't know what the data set is.

DR. CHARACHE: So, this is the summary, then, of 17 patients, 11 of whom it agreed and six of whom it did not agree. Is that the right number?

DR. SPECTER: Correct. If one looks at the data, it says that 11 additional specimens were picked up. I don't know how many patients that represents. But six were not. There are no false negatives in that data set.

The question is, is that a good data set to use or not, is what I am asking, and does it tell us something important about the value of doing smear-negative specimens.

It certainly looks like it shows benefit, but I

don't know if it is a reasonable data set. I would be happy if the statisticians would comment on it. I certainly don't have the expertise.

MR. DAWSON: I had that chart included only to make a particular point about my concern about fresh and frozen.

This was the definition of the data set, based upon the questions that went back to the company in January. It went further beyond that, and I really have to apologize because a lot of this took place in the review process.

This data set only looks at the frozen MTDs, and we felt that it would be still more realistic if we would allow the inclusion of the second MTD. So, this was kind of one step back, as far as we were concerned.

DR. CHARACHE: Do we have the information on what happens when you include the second data set?

MR. DAWSON: We do, but that is only the fresh and frozen combined. That was for a request that I made.

DR. RELLER: This analysis, was there -- the fresh versus the frozen for smear positives, there wasn't much difference, in smear negatives there wasn't much difference.

Then, when you did it somehow a different way,

with smear positives and smear negatives, the smear negatives were possibly different. I mean, what was the purpose of this analysis?

MR. DAWSON: I realize this gives rise to confusion. We were trying to find a data set that, to us, corresponded to intended use.

Given that the intended use is still being debated, we developed a sort of definite idea of what we think it is.

So, we wanted to use that part of the data, with that idea of the intended use. So, unfortunately, it has a narrow focus, but we think it is more appropriate for that purpose.

DR. RELLER: What do those P values mean?

MR. DAWSON: The P value at the lower left corner is for the McNemar test, and that was just a comparison of the clinical status and the MTD2.

Essentially, what we compared there were the 11 positives by MTD2 versus 17 by clinical status. I think those numbers are the same but for standard variation, and the low P value means that they are not included.

DR. RELLER: So, as we come across these bottom four cell blocks, the first one, smear-positive, P value of one. Smear negative P value is one. Then we come

across the smear positive, the 3.7, which I presume there is some relationship between the first block of smear positives and the second block, and the first smear negative and the second smear negative. What exactly is it, and where did those numbers come from?

MR. DAWSON: The implied conclusion is that maybe smear-positives is not so much difference, whatever went into those different blocks, but somehow, the smear negatives, by this analysis, there may be a difference of a suspicion of a difference. I mean, the P value is less than .05.

DR. RELLER: Just. What are you really trying to get at here, and what does it mean.

MR. DAWSON: I think what those four blocks mean down at the bottom of the chart is that everything looks okay except smear negative in fresh specimens. That is basically why I included this.

This is my best cut at -- or was as of January -- my best cut at the data. I had a concern that smear negative was rather different in the fresh specimens than in the frozen, because everything else looked fine.

The smear-positive results are fine, and the smear-negative in frozen specimens. I sort of had the view that testing worked better if it was frozen.

DR. RELLER: That is exactly why I am asking these questions. Now, we come to this smear negative fresh only. Is that excluding the bad primers?

MS. JONAS: Yes.

DR. RELLER: So, that is without the bad primers. Okay. So, we have the fresh specimens with smear-negative. That is what we are interested in, smear-negative fresh.

Statistically, it looks like it is less good performance than frozen specimen and smear negative? Better performance? Better performance on the fresh, than are frozen.

MR. DAWSON: Than the pooled data.

DR. RELLER: Just reinforcing what the --

DR. GUTMAN: Can I try to clarify this? The intention of this was an exercise in understanding poolability.

You are absolutely right, that if you decided that poolability was taken off the table, either for biological or statistical reasons, then you are looking at the most pristine data set, that this is the first sample against multiple cultures. So, it may be small, but it is the most pristine data set.

DR. RELLER: So, it is a feeling that this is a

really good test, if we had the data to conclude that. It is just that the numbers aren't there. It is sort of unfortunate, because of the glitch in the primers, that the numbers aren't there.

It is the fresh done right away, which in terms of the old flow of things, the real value of this test, either smear positives or smear negatives, is how good it is now.

Then there comes the question, for me at least, of are the numbers, with the way it should be done now, with a fresh specimen, adequate, or at this point adequate, in terms of numbers to be included in the label, enough numbers to be convinced that it is effective for the purpose that it is intended to be used for.

It is not that it won't be or isn't, but do we have the data base to conclude that.

DR. SPECTER: We have one set of data that looks pretty pristine and we just looked at that, and it looks pretty good. The question is, are the numbers large enough to make that conclusion.

I think, drawing into that, though, you have to say that it is hard to generate those numbers to begin with. I don't want to put numbers in the FDA's mouth,

but at least they were at least neutral on the sufficiency of the data going forward.

Again, with the idea of least burdensome, is that from a regulatory point of view, is that enough data, in terms of safety and efficacy standards, to say that the test works in the normal situation.

I think the other thing, from the more practical point of view, is -- as I said before, there are not that many tests marketed. The sponsors have to include more data and it may more worthwhile to stay with what you began with. It may not be practical to gather that much more data going forward.

You have to make a decision as to whether the test is going to be developed after.

DR. CHARACHE: The only thing we see from the data set thus far, is the very nice looking data, that shows the comparison with the clinical definition.

The clinical definition also was challenged. We might want to consider whether the numbers are accurate, whether more comparisons should be looked at from that set, that we have not had presented today or whether we have all the information that is needed today.

DR. EDELSTEIN: I may have a way out of this conundrum. The data from the fresh specimen set actually

looks pretty good.

The sensitivity is a little bit higher than the pooled data. The specificity is considerably greater. The question is, if we look at the pooled data, what we are saying is, that may be the worst case scenario for the performance of the test.

If we say that, based on that worst case scenario, that that looks okay in terms of safety and efficacy, then, while we wouldn't address the issue of what the true performance of the assay is, we will have a pretty good idea of the worst case situation.

For me, I think we need to move beyond this. In answering the questions about poolability, I was answering the specific question raised, but the issue is, accepting that the pooled data set may not be perfect, overall, its performance is pretty much comparable to the fresh alone. In fact, the major changes are that specificity is a little bit lower with the pooled data set than with the fresh alone.

DR. CHARACHE: I think it is probably a good time to take a break and catch our breath.

DR. SANDERS: I have a comment before we do that, because it tags onto what he was saying, and it has to do with this whole issue of ultimate recommendation.

I have a reaction.

This may be one of those areas where it is ultimately a condition of approval, certain conditions for a post-marketing data base. I just throw that out.

DR. CHARACHE: Brief break, 10 minutes, really, 10 minutes.

[Brief recess.]

DR. CHARACHE: I would like to begin by reminding the panel and colleagues responding to questions, it is important to use the microphone.

The panel members can ask questions that the manufacturers or other consultants can respond to, but we do have to use the microphones when we do it. We can't just talk across the chairs, unfortunately.

MS. POOLE: One addition, if the audience wants to put a question, could you wait until the chair recognizes you before you start to comment?

I would also remind the people in the audience, if you would take everything that you came with, and any trash, if you could place it in the garbage containers outside, thank you.

DR. CHARACHE: To return to the question that we were discussing, the question had been raised as to whether the number of smear-negative TB cases, with the

analysis that we currently have available, which is adequate to make a decision at this time.

I think there was agreement that the specificity of the first sample data looked very good, when compared to the clinical diagnosis from the pooled data.

We have to continue to consider whether that one piece of information provides enough information, or whether there are additional questions that we might have, about the analysis of the particular patient population that we are concerned with; namely, that associated with the extension of the indications for use; that is, the smear-negative, TB-positive group.

That is associated, whether it is -- perhaps we can address first question B. Is the data that we have available, or provided prior to this session, adequate to characterize individual site performance for the use of this device. The pooled data, obviously, look excellent.

DR. HAMMERSCHLAG: I think as I mentioned before, it is just that it is a very heterogeneous group of sites, with some of them having very small numbers, and Zurich, where you have 13 specimens, and 10 of the patients are smear-positive, and our limited experience at a site where there are 50 patients and two to three percent smear positive.

It is hard to -- it would be difficult to extrapolate from the performance of one to the performance at another, just because of that, even though in aggregate you can interpret the data as very good.

DR. GATES: Isn't that the idea? I mean, if we are looking at the data in the aggregate and sampling across a lot of different sites in a lot of different populations, it seems to me that gives it more balance.

DR. CHARACHE: But we are addressing question B. Can we characterize individual site performance from these data.

DR. SPECTER: I really don't see how we can answer that question based on our previous discussion, because all the data we have been presented is fresh and frozen together.

If we won't accept fresh and frozen together, we don't have any data to evaluate and answer that question.

DR. CHARACHE: Any divergent opinion on that particular point? If not, what types of data or further analysis should be used for laboratory site performance evaluation?

Here we probably can be helpful, because the data may already be there, and we are just saying what it is should be shown. What else would people want to see,

to be able to evaluate individual sites?

MR. REYNOLDS: Is it even necessary to have individual sites?

DR. CHARACHE: I think the purpose of individual sites is to show consistency. You can't have all the data from one location, which may not apply when a different laboratory sets up the same assay.

The question really is, what data would we want to request be provided, in order to answer that question?

DR. O'BRIEN: In discussing this question, can we consider data that aren't a formal part of this study that have been presented to us?

For instance, the study from Galveston, that was provided to us from FDA, does have data that are relevant. There was another paper included in the Gen-Probe submission, that has data relevant to this, both from single-site large studies.

DR. CHARACHE: If it was presented to us, yes. It should be fresh without frozen for individual sites. Would you -- is that fresh but not frozen for individual sites, the papers? Which reference should we look at?

MR. FRIEBERG: Would you like it? We have it summarized.

DR. SANDERS: We have the Galveston study in our

packets. It is just that we have several volumes for this particular project. I did see it.

DR. EDELSTEIN: I don't understand this. How can we use the data from another study to characterize this evaluation?

DR. CHARACHE: It is another study. It is not this presentation.

DR. O'BRIEN: We may be able to use these data to characterize the overall performance of the assay. Although we have a published paper to read, we don't have all the details of the exact conduct of the study.

I am sure it was impeccable, but it would be so unlike any other submission that we review, where we would see all the nuts and bolts.

DR. CHARACHE: I think that is right. I think there have been questions raised here in terms of using the clinical definition when each study site has its own definition of clinical tuberculosis.

I must confess, I was questioning, of the smear-negative group, which we are really concerned about, what was the distribution of the number of cases that went to the panel, and to what extent did the knowledge that was a positive culture impact on the definition of whether it was TB or not.

These are all things that we would know if this was analyzed as a separate subset, but we don't have that data right now.

That is my only comment. I would like to see other thoughts.

DR. O'BRIEN: I thought the patients presented to the panel were well characterized. No?

DR. CHARACHE: We don't know if they were smear positive or smear negative. If they were disputed, and they went toward the smear negatives, it would impact on the assessment.

DR. O'BRIEN: They were virtually all smear negative, I think.

DR. CHARACHE: That went in the panel. It has the potential to impact on the definition of a clinical diagnosis, if the predominance of smear negatives went to the panel and not an equal number of other studies.

DR. O'BRIEN: All of them were culture negative as well.

DR. CATANZARO: No, that is not true; I am sorry.

DR. O'BRIEN: The question is smear negative.

DR. CATANZARO: That is not true either.

DR. CHARACHE: I am sorry, could you please?

DR. CATANZARO: I would be happy to. The criteria for going to panel is that they didn't fit the two criteria that were described.

One criteria is they had a diagnosis of tuberculosis, and they had two positive cultures -- not one, but two -- and a clinical syndrome compatible with active tuberculosis.

The criteria for having tuberculosis excluded that all cultures were negative, and there was no clinical syndrome that was in any way close to compatible with tuberculosis.

So, anyone who didn't fit those criteria went to panel, and there were 40 of those people. Some of them were, in fact, culture negative.

It turned out that, for the whole study group, there were five culture negative patients. So, that is really not an important issue.

There were some smear negatives I those as well, and I don't recall the number, but I do know that they were not all smear negative.

In fact, there were people who had one smear positive and one culture positive, and they would have gone to the panel a priori, and there were a number of those, quite a number of those.

DR. GUTMAN: I don't think this information -- I think this analysis was probably not in the package.

DR. CHARACHE: I didn't see it, because I was looking for all the data I could find on the smear negative TB.

DR. CATANZARO: The FDA was not impressed with the function of the panel, and a lot of the panel activity was not reported because the FDA didn't feel it was relevant or something. I am not sure of the reasons, but the FDA chose not to accept it. The panel, for me, was critically important.

DR. CHARACHE: Our question here is, what information would the panel find useful in addressing the differences between, or the similarities or differences between the study sites. Would anyone like to come up with a list?

DR. SPECTER: I think the first one would be to analyze it strictly on the basis of fresh specimens.

DR. CHARACHE: So, we would like to see it on the basis of fresh specimens and we want to see the smear negatives, diagnosed as positives. We want to see that data for each site?

DR. SPECTER: Yes.

DR. CHARACHE: What else do we want to see from

the site? I have expressed an interest in knowing how many went to the panel compared to other groups of patients, and how that was resolved, whether it was just how the culture compared, or as well as how the diagnosis compared.

Dr. WEINSTEIN: I am putting this out as a question, not really as a suggestion. Does anybody believe that it would be useful to look at the data from those sites that enrolled many patients -- 50, 58, I can't remember what all the numbers were -- and look at those data and then look separately at the data from the sites that enrolled 13 patients, 15 patients.

If there is a learning curve and that sort of thing, is that going to be a useful exercise. I ask it as a question.

DR. CHARACHE: Would you want that or in any case, whether there were discrepancies between the large pools and so on? You will see each center. Any other thoughts on that question? Can we go to question two?

The applicant suggests the following interpretation of the results.

DR. GUTMAN: If you look on the printed copy of the questions in the handout, it actually has the table on it.

DR. CHARACHE: This points out that the interpretation results currently recommend MTD-positive, MTb-RNA detected as the interpretation, as meaning positive for MTb or MTb plus MOTT, regardless of smear status.

Then in brackets it says, definitive laboratory criteria for TB diagnosis. I don't know what the brackets mean. Can someone tell me why there are brackets?

MS. SHIVELY: In Dr. Simone's presentation, she indicated that case definitions recommended by CDC indicate that a nucleic application amplification test for smear-positive patients may be used as laboratory evidence of MTb for confirming the case of TB.

DR. CHARACHE: So, this statement that is in brackets would indicate that this would be a case of tuberculosis that is smear positive and nucleic acid diagnostic positive.

For MTD negative, MTb RNA not detected, interpretation, MTb not detected or low numbers of MTb with or without MOTT, or MTb present, not detectable due to inhibitors. Those would be the three interpretations.

Then it continues, MTD testing of another specimen recommended, if TB clinically suspected. That

recommendation is made regardless of the smear status.

The question A is, should there be instructions that separate the expected performance for smear negative versus smear positive specimens.

I think this gets into the point four, which was raised by Mike Wilson, which is the question of prevalence, which might impact on the significance of the interpretation.

There might be other factors that bring in other points. I am trying to make sure that we address all the points. Comments by the panel, please?

DR. EDELSTEIN: I am always slow to make a comment. I think that there definitely have to be different criteria for smear-negative versus smear-positive patients, interpretive criteria.

Just taking the first example of smear negative nucleic acid amplification positive, and even accepting the positive predictive value of around 80 percent with a confidence interval of around 50 to 96 percent or so.

To me, that has a different meaning than a positive predictive value of 100 percent, which would be the case for smear positive.

I am not exactly sure what the interpretation should be, but it should probably say something about,

interpretation of the results of this in a smear-negative patient depends on pretest probability of disease.

It may be helpful, in the product insert, to give a table listing positive predictive values versus pretest probability in the smear-negative population.

DR. CHARACHE: If we stay with that for a moment, what population would you base your predictive value on, your incidence?

Would it be those physicians who thought it likely enough that it could be TB that they sent a sample to the TB lab?

Would it be those patients put on isolation for TB, with that degree of certainty? Would it be anybody who is HIV positive with TB? What kind of guidance could we provide?

DR. EDELSTEIN: I am not sure I understand your question. You are asking in terms of indications for performing the testing?

DR. CHARACHE: Interpretation.

DR. EDELSTEIN: That is going to be nearly impossible for the laboratory to do, and I think that is something that the clinician has to do. He has to estimate the pre-test probability.

There can be guidance given in the product

insert for resources to estimate pretest probability, but I think it would be real difficult, I think, for individual laboratories to do that.

Even if we know the overall pretest probability in our total population, that wouldn't define the pretest probability for the individual patient.

DR. SANDERS: In the proposed package insert that they have provided us, on page 32 of book number three, there is a section that has test interpretation.

It doesn't exactly state what is in our questions from FDA.

I am perfectly satisfied with what I read there, and particularly with the sentence that says, the MTD test should be interpreted in conjunction with other laboratory and clinical data available to the commission.

Based upon the level of clinical suspicion, testing by additional assessments should be considered. It lists the IRB cut offs, the clinical range, I was fine with that.

In addition, on page 37 of the proposed package insert, at the top table, it gives, at least me for the clinician, relevant information regarding smear-negative patients and how to compare the MTD with BacTec, other cultures, and it gives the sensitivity and specificity,

positive predictive value.

So, it gives all the information right there, in a way that I, as a clinician, could easily understand it.

I think question number two, as outlined in our FDA handout, actually makes the issue more complex.

DR. CATANZARO: I was going to say almost the same thing. I think the criteria for the study are what you should use as the criteria for testing.

The criteria for the study, where people were suspected to have tuberculosis, as we saw in the data presented by Roxanne, 80 percent were in isolation, and there were other criteria that were specified there as well.

A smear status was not a criteria for the study, and that is a post-facto fractionation, which I think is entirely inappropriate.

MR. REYNOLDS: I don't understand that comment. We are talking about interpretation.

DR. CHARACHE: If I may, I don't want to get into a dialogue like this, because we have a lot to cover.

Perhaps I could ask again, how one could perhaps provide some guidance. I think Paul has called our attention to the fact that you can do it by populations,

but it wouldn't necessarily apply to a single individual, and I think that is a challenge.

I did, since we just had an infection control meeting on this subject, look at the prevalence of disease in patients who were put into isolation in our institution over a five-year period.

The prevalence was as high as six percent positives, of those who were on isolation, if you assume that everyone who was positive by culture was put on isolation, which they were not.

That is the very highest end, to as low as .7, before we restricted who could be put on isolation.

If you look at the whole curve of the positive predictive value, that means we were somewhere between 20 and two or three false positives for every true positive in our institution. That is an inner city hospital with a very large AIDS population of 3,000 or 4,000 patients.

I think it is very clear that our institution would have been compromised. We would have had approximately, all together, if we had done it with the total population, perhaps as many as 75 or 100 people called TB who were not, because of the false positives in that particular, very low prevalence group.

So, that is one of my reasons for being very

concerned about the positive predictive value which you have heard about a couple of times this afternoon.

We need to look at that, as well as, or in balance, perhaps on a population basis, with this issue of the very clear advantages of making an early diagnosis in this population.

MR. REYNOLDS: I think I agree with Dr. Sanders.

I don't know how you can put a specific statement in there. As long as you put a general statement in there with positive and negative predictive value, varying with location, and you have to appraise yourself of what the incidence is in your facility and then you that information accordingly, I think that is about the only way that you can do it, is to put some sort of general statement like that in there.

DR. HAMMERSCHLAG: Looking at it another way, and oversimplifying it, the reality is, there is no such thing as a test that is ever going to be 100 percent sensitive for a specific set of circumstances.

The question is, what can we tolerate and what can we consider to be adequate. I mean, extrapolating from experience with, say, chlamydia trachomatis diagnostics, the Centers for Disease Control have come up with some standards for sensitivity and specificities.

I can't give you the exact numbers, except that the test is in the high 90s somewhere, and if a test doesn't perform according to that, they will not recommend it.

The question is, there is always the possibility that it could be a false positive. That may be greater or lower depending upon certain characteristics that the population may have. You have to just always take that into account.

This may sound rather simplistic, but some people just don't get it. It must be true; the lab told us; it must be true.

I always tell my students, 100 percent never happens. It is just that I think it is a little too positive.

There needs to be some kind of qualification, that they have to know that there is a probability that it is a false positive, although the specificity may be in the 90s somewhere.

They have to take that into account, knowing their population characteristics, and it is a question of what they can tolerate and how to get it.

DR. GATES: I would just like to concur with what Dr. Sanders was saying, in the sense that the people

writing these sorts of tests are, from a laboratory point of view, pretty sophisticated.

You would think that they know what prevalence meant and what effect it had on positive and negative predictive value, and that you never absolutely address each lab's prevalence. You depend on the lab to know what it is and conduct the test accordingly.

DR. WEINSTEIN: I agree that the leadership in many laboratories is sophisticated, but I would argue that the leadership, in many laboratories, is unsophisticated, particularly if you get to hospitals where you don't have doctoral level leadership in the laboratory, where you have, with all due respect, pathologists who are directing microbiology laboratories who had a total three months of microbiology during their five years of training, and are sort of doing this as an aside, while they make their money on surgical pathology.

I do think that there is a potential problem there. I think what might be useful would be, in the package inserts, to include the graph that is designated B-1, book number three, page 87, that would at least give some guidance to people who would then know that if they had low prevalence in their particular institutions, that they have to be aware of the possibility of false

positives.

DR. CHARACHE: May I point out a generic question? What percentage of false positives would the panel think means that institution shouldn't do the test?

What is your tolerance level? Fifty percent? Ninety percent false positives? That is really what it is going to come down to.

Should there be institutions who don't offer it to their patient population because the likelihood of true positives is too low? What percentage false positives -- should that be a determinant? Does it reflect the sensitivity?

What does the group suggest is tolerable in terms of false positive rates? Should we be worried about it?

DR. GATES: Is that within our purview, to make recommendations to clinical sites in terms of their false positives level?

DR. CHARACHE: I am not saying that that necessarily has to be, but it might provide guidance if one were to say, if your incidence of positives is below a certain level.

DR. HAMMERSCHLAG: I think that may be more in the reign of the Centers for Disease Control and good

practice guidelines and testing. That is usually where the recommendations come from.

That is sort of an independent issue from this warning, saying that there are going to be variations in the way this test performs based on population and prevalence, and that you should never accept it just out there at face value.

I agree that the leadership in laboratories varies. The ASM had this editorial a couple of months ago about dumbing down in the microbiology labs.

Many places, where they used to have doctorate level people, now these tests are being done by people with bachelors degrees or associate degrees.

There is a real possibility, of course, once the test is out there, it may not always be at higher level laboratories. It will go into commercial laboratories and small community hospitals.

Just judging from experience, again, that I had with inappropriate use of chlamydia tests, it is really wild, some of the stuff that goes on.

DR. RELLER: Don't the boundaries on guidance for smear-negative and smear-positive specimens depend on whether the recommendation is that one should do the test on smear-negative as well as with already-approved smear-

positive specimens?

One of the questions that came up with the fresh specimens, for example, versus the frozen one, is that there seemed to be, although the numbers were small, differences in sensitivity.

The caveats that one would place on the test have to do -- interpretive caveats -- would have to do with sensitivity and specificity.

I don't think I know what that is for smear-negative specimens, yet. There is confusion about use. Just to amplify Dr. Hammerschlag's comments, in the presentation from the CDC, when the impact of active NAA on TB, on page 86 of the slides, smear-negative NAA positive -- this is how one would use this, smear-negative, NAA positive; i.e., those with potentially less infectious TB.

The impact, reduced delays in the initiation of therapy for patients with smear-negative TB when there is a low clinical suspicion.

Well, I thought those were the very patients -- I mean, I think one has to see what you have before you describe how to interpret it.

The question I would like to pose, coming back to the data base that we have under consideration, that I

think there is general agreement that we are most interested in, but what we are most interested in is how this test performs on smear-negative fresh specimens.

How many of those 17 patients, that are in that best data base -- best data base by most people -- how many of those people were, in fact, culture positive? Do we know that?

I think, coming back to Dr. Wilson's comments at the very outset, I don't think anyone would continue therapy ultimately in someone with non-confirmed disease, specifically if they ran into difficulty.

I applaud trying to get a really good handle on someone who actually had tuberculosis. From a clinical standpoint, there are criteria that one -- this is what is so great about tuberculosis.

What is so difficult about defining high suspicion or whatever is that it is a sneaky disease. There is a wide range of manifestations.

How do you define someone with syphilis? Absolutely. In the end, you have to confirm it with a laboratory test.

Someone who had pulmonary infiltrates that was at high risk, one could add in all the pretest probabilities, who had one positive culture, I don't

think anybody would say they didn't have tuberculosis.

I am more interested, ultimately -- I think most patients who complete a course of therapy, in fact, there is no reason, with good laboratory techniques, that most of them who end up completing a course of therapy, in fact, do have culture confirmation.

I think in the end, for the purpose of getting some objective boundaries around it, we have to consider the culture-positive patients.

That is why I want to know, out of the 17 in the pristine data base, how many were culture positive.

MR. FRIEBERG: I think in the April 2 submission, page 55, there is a table that shows culture positive and smear negative cases.

DR. CHARACHE: What page?

MR. FRIEBERG: Page 55, book 3, April 2. Hopefully that is what you are looking for. I think the answer is 17, all 17 of those were culture positive.

DR. CHARACHE: So, that suggests that perhaps the resolution was because they were culture positive, which is part of the decision.

DR. RELLER: I wanted to actually know how many culture confirmed patients that we have, in which to assess the performance of the latest version of the MTD

for smear-negative patients.

I think when we know that, then if we are comfortable with those data, then one can, based on that sensitivity and specificity, put the boundaries around this.

The prevalence is an important consideration and it serves an educational purpose. For the individual patient in whom one is considering the diagnosis of tuberculosis, one could have a very high pretest probability of tuberculosis, based on clinical and epidemiological grounds, in an individual patient in a low prevalence area.

It is the individual patient that is critical in making the diagnosis of tuberculosis. I think, wherever we end up, at whatever time -- today or next year -- that the issues of the pretest probability and what the actual performance on culture-positive, smear-negative patients is, is the kind of descriptive language that would go into the package insert to help guide the appropriate interpretation and use of the test.

DR. CHARACHE: I wonder if there is further information, if all 17 individuals that were considered to have tuberculosis, were culture positive, whether that was the basis for deciding that they had TB and how that

corresponded to the population as a whole.

I am interested in that, because that group had no discrepant results. I am just wondering whether there are similar smear negatives that were MTD positives, that followed the same clinical appearance with the exception of a positive culture.

DR. CATANZARO: I don't have the answer to that question, but I wanted to implore that you remember that the rest of the cases -- not the 17 that were culture positive -- but the ones that didn't have TB, but that if they had gotten the information, would not be like the patient who has positive X-rays, smear negatives, bronchoscopy negative, but the clinician felt the patient probably had TB.

MTD would differentiate those patients, would not cause people to waste tons of money on people that are smear negative and MTD negative.

I am just suggesting that you remember both sides of the coin. Focus on the 17, forget about the 200.

DR. CHARACHE: A point of clarification. We know that in the study population as a whole, there were some that were decided to be false positive.

Is it the recommendation that if you have a

positive MTD and a negative smear, that all those patients should be treated for TB? Should they have investigations in introduced by the public health department?

DR. CATANZARO: My personal recommendation is they should have another test.

DR. TUAZON: I think the bottom line there is their clinical judgement. If you think clinically the patient is highly suspicious of having tuberculosis, you would ignore the MTD and the smears.

Let's return to the question. Should the instructions for used in the labeling include information to clarify differences in expected performance for smear negative versus smear positive specimens.

If there should be differences between the two, then let's sharpen what we think they ought to be.

DR. SPECTER: My feeling is that, if you look at the table that we have been provided, it says, MTD testing in another specimen, TB clinically suspected, regardless of smear status.

I think it has already been suggested, besides that, that if you have a negative, it should be repeated as well.

This recommendation is that if a definitive test

is received with the MTD, then it should be repeated. Then I think you can consider smear positive and smear negative the same.

If that recommendation to repeat is not going to be followed for all of these tests, then there should be distinct recommendations for smear positives versus smear negatives.

If you look at this line and it is not clear where it is placed on this table, whether that just refers to the MTD negatives or to everything, we need to clarify that point.

DR. GUTMAN: I can clarify the question. We are focused on the negatives, not on the positives, and one of the intended uses is that there would be different interpretive criteria.

I think they would probably be different among the group we have reviewed, that they probably meet the global interpretation of the guidelines for suspicion.

Many of us in the review group sort of think, whatever the instructions are, that for these sorts of things, the data seems to suggest difference in performance in those two subsets -- smear negative versus smear positive -- you are putting out the data, to partition the data into smear-positive versus smear-

negative performance, or are we, in fact, raising a valid question that that is wrong and they should be required to put different performances in partitions.

Obviously, their instructions may or may not be the same, depending on what that performance shows.

DR. CHARACHE: Comments? I think what I have heard Dr. Specter say is that he feels that the recommendation for a repeat test should be clear, as applied.

How far would you want to go on that? Would you want to say must be, in order to interpret? Would you want to say should be or maybe?

DR. SPECTER: Based on the data I have heard presented here today, I would say it must be.

DR. CHARACHE: Comments or thoughts, Dr. O'Brien?

DR. O'BRIEN: I am not sure we have enough information on repeat testing to make that judgement, particularly if patients have a low probability or a low clinical suspicion of tuberculosis.

Maybe they shouldn't have this test in the first place, but I am not sure what value repeat testing would be in those.

Those with a higher suspicion, maybe it would be

reasonable, but again, do we have the data presented in the study to make that determination?

DR. CHARACHE: What is the consensus? Do we have that data that says what our recommendation should be on repeat testing of smear negatives?

DR. O'BRIEN: We were presented some information demonstrating that the sensitivity increases a bit with second and third tests, because it does work with smear and, particularly, culture.

DR. CHARACHE: That was mixed specimens?

DR. O'BRIEN: That was mixed specimens, right.

DR. EDELSTEIN: As I see it, the problem with repeating the test is, in the particular situation of a smear-negative patient, if the initial test is positive, the interpretation is going to be based on the pretest probability of disease.

How are we going to suggest, if the repeat test result is divergent, to do a retest. So, it raises a lot of issues that we can't address from a laboratory standpoint, that have to be based on what the clinical suspicion of disease is.

DR. RELLER: The best use of the direct test is to get it done on a good fresh specimen, when initial decisions are being made.

When the test is positive, it needs to have a culture and a confirmation and susceptibility testing. When it is negative, it certainly needs to have a culture, because the test is not appropriately sensitive.

If the recommendation for a package insert is that, if the test is negative, that one needs to do a culture for smear positives, it is unequivocally necessary, with lesser sensitivity, to do a culture on the smear negatives.

I mean, you have to do further testing anyway, to confirm, because of issues of specificity, or to exclude, owing to issues of sensitivity.

DR. CHARACHE: So, you are emphasizing the necessity -- that what must be done is not to repeat. What must be done is the culture.

DR. RELLER: I mean, I am sorry, you have to wait for the ultimate answer. I mean, this helps in the specificity of the test in smear positive, and the specificity in smear-negative patients, is the real value, applied clinically in a patient with a high pretest probability based on clinical and epidemiological grounds.

DR. CHARACHE: Did you have something to add?

DR. SPECTER: I just wanted to make it clear

that my comments earlier did not preclude that. I particularly agree.

I think one has to take into account both that culture must be done, and that clinical evaluation is going to be a critical factor. I take those as givens.

I still think you need confirmation of your initial results in order to make the same judgement in using this test.

DR. CHARACHE: Let's take one sweep around the panel and ask if there is anything else. We heard that there should be information presented relevant to the difference in results, different predictive value, based on the prevalence in a given population.

There has been a sense that it would be very difficult to specify, other than with clinical restrictors, how that information might be employed in a given area.

We heard that one must do a culture, not that they should.

DR. HAMMERSCHLAG: I generally concur. I just have to give a little more background information. People can interpret the results -- and remember, that the results are not 100 percent absolute in either direction. They have to use them appropriately. Of

course, they have to repeat a specimen.

I think that should be included. I don't think it is going to be overwhelming, or add that much more to the insert.

I think one variation might be that graph, just to show that, again, it is not absolute.

DR. SANDERS: In the old and the proposed package insert, there is a statement there that says, the MTD test should always be performed in conjunction with mycobacterium culture.

DR. CHARACHE: The question is whether it should be should or must. There is a very big difference.

DR. SANDERS: I would say it should be a must. I think that there would be an enhancement of this whole issue of AFB smear-negative patients if the prevalence table were also included later on.

DR. WEINSTEIN: I concur.

MR. REYNOLDS: I agree with the comment about the graph, putting the graph in there. Insofar as saying must, right now, how often do you have the doctor order an AFB smear without a culture?

DR. GATES: I think the graph is a good idea. I also think you shouldn't get too prescriptive. I go along with should.

DR. RELLER: I was trying to think of the right word. I think it must be a must, and that, just parenthetically, the kind of lab that we would all like to have our work done in, would not do a smear without getting a culture.

What do you do with a negative? What do you do with a positive?

The importance of this language is that if there is a crack -- in a way, should is a crack. In today's environment, HUM-Vs would be driven through a crack, if it was economically beneficial.

DR. SPECTER: I am in favor of must and inclusion of the graph.

DR. O'BRIEN: I don't have strong feelings about must or should, but very definitely not only the graph, but I would put a warning to indicate that the positive predictive value given in the table for smear negatives of 75 percent was based on 11 percent prevalence of disease in that population, which we won't find anywhere in any laboratory doing this test, and that the positive predictive value is likely to be much lower, and that needs to be taken into serious consideration.

DR. EDELSTEIN: I vote for should. I think that there should be a clear statement regarding the

interpretation of the test, to take into regard pretest probability, and also I concur that there should be a statement specifying what the prevalence of disease was that was studied, and saying that the performance of the test may not be the same in your patient population.

DR. CHARACHE: We will complete, and then I am going to ask one question relevant to that.

DR. TUAZON: I don't have any strong feelings about must or should, because you are doing it anyway. I concur with the other recommendations.

DR. WILSON: I vote for must. I agree with Dr. O'Brien. I think there needs to be a fairly strong warning.

In many parts of the country, regardless of the individual patients, the prevalence is so low, that the positive predictive value could well be very low for this test.

I think that people performing this need to know that and need to be able to communicate that.

DR. CHARACHE: I agree with must. I am going to raise an additional question pertaining to prevalence and I will give you one other piece of information.

I told you what the prevalence was in our inner city hospital for patients on isolation. Last year, for

which we have the full data, which was 1998, the prevalence of positive culture, of those received in the laboratory is 0.2, which is a very long way from 11 and I don't even think it is on the curve, and we are a very large, inner city hospital, with an AIDS population.

My question is, given the fact that everyone agrees that you can't draw any conclusions from prevalence because it is so individual, should the prevalence reported in this study be put in a package insert? It says it will not apply to any of the hospitals using it.

DR. HAMMERSCHLAG: I think actually somebody suggested that. I think it is probably late and my neurons aren't functioning as they should be right now.

I also agree, again, I agree with Dr. Reller about the crack, and why we have to, again, emphasize the word must.

If this goes into some -- it is very possible, that even if you say that the test should be run by laboratories who have the capability of doing culture, I almost think that almost anybody would be able to buy it in a community hospital.

You can't just say, we will not sell it to Franklin General Hospital in Nassau County.

It doesn't mean that it won't be done in clinical chemistry, and that doesn't mean that it won't be done in a laboratory that isn't going to be doing microbacterial cultures.

MR. DUNN: It is a complex test, so theoretically the laboratory has to be able to do it.

DR. HAMMERSCHLAG: There are some that are. It is hard to say. There are certainly other tests -- they may even get involved in other amplification test. I just don't see how you are going to be totally able to tell them to restrict it to hospitals that only have the capability of doing mycobacterial culture.

DR. CHARACHE: Does anyone else want to comment on whether the prevalence results that were found in the clinical trial should be published as part of the package insert?

DR. WEINSTEIN: I agree with you.

DR. RELLER: That means that they should or should not be?

DR. WEINSTEIN: That they should be published in the package insert.

DR. CHARACHE: The prevalence? No, we are asking whether it should be shown to have a positive predictive value of whatever percent, based on the study

that has a prevalence of tuberculosis of whatever percent.

DR. HAMMERSCHLAG: Sure, and then you can say that the performance may or may not be the same, and then follow up with the prevalences in the population, and you can consult with the graph to get an idea of what it might be.

DR. RELLER: That is the point.

DR. CHARACHE: Could you clarify the point that you want to make?

DR. RELLER: There is no use of the prevalence unless there is some educational construct in the package insert that relates that the expected positive and predictive value varies with the prevalence or the pretest probability.

To put in one without the other is misleading. The performance in a lower prevalence or lower pretest probability will not be the same.

The prevalence doesn't make any sense unless you link it with what effect that has on the probability of a positive meaning, or a negative meaning. I think they are linked together.

Then one can position based on -- it affords a couple of things with having some latitude in this. One

is the reality that, if you are going on a fishing expedition, you are liable to end up with a red herring.

You need to have, you know, a reasonable clinic.

Second, you are most often going to have those in a place that has high prevalence.

Once you get the understanding down, then you have laid the ground work for the kind of communication that I think everyone here wants to see, between the laboratory and the clinician, to say, you know, go to the lab and do this test. It could be useful.

Somebody who is picking it off to exclude TB in a low prevalence population, they are deluding themselves.

DR. CHARACHE: We are going to have to move around very quickly. We have one hour left in the building. We have got more questions and we haven't voted. Any more thoughts on how to add any other guidance or caveats to the label, appropriate to ensure safe and effective use of the MTD with smear-negative specimens?

DR. O'BRIEN: In answer to your question, I think if we are going to publish data, or suggest that data from the study be included in the package insert, then this should be presented.

I don't know how much guidance we can provide to indicate that, if this test is used in smear negative indiscriminately, in most settings in the United States it will have a very low positive predictive value.

If used in a patient who is suspected of tuberculosis, with a suspected a priori likelihood of TB of around 10 percent, it will have this predictive value based on the study data.

DR. EDELSTEIN: I agree with that. What I would like to see in addition -- I hope this doesn't sound too heretical -- is the 95 percent confidence intervals of those estimates.

I would be willing to accept the confidence intervals for the combined data sets, because I think that is the worst case scenario.

DR. WILSON: I agree with what has been said. That caveat needs to be in there. Even though the idea of using those tests is for a particular group of physicians working with particular patients, the reality is that you have an enormous number of specimens coming in from less qualified physicians who don't routinely care for patients with tuberculosis.

They are the ones who really need to understand that prevalence in that community greatly affects the way

the test result is interpreted.

So, the test won't be used in idealized situations. I think that makes it even more important that that caveat is in there.

DR. CHARACHE: Could we see the third question?

DR. GUTMAN: I hate to be so repetitive. I know it is late. Are we under the assumption that you are thinking of this as a separate part of the package insert, from these issues that deal with the smear negative and smear positive, that you are talking about different prevalences?

We are certainly appreciative of the insight, in terms of the notion that prevalence might be useful information on the label.

We also think there is a spectrum bias here, in that there are different sensitivities and specificities between the two subsets.

We are asking you to either support that notion or to suggest that notion is wrong. Somehow or other, that might have gotten missed in the conversation.

It is okay to say we are wrong, but we would be surprised and we would go back. If the performance of the two subsets is different, it is calling for some kind of differential labeling, maybe.

DR. CHARACHE: Let's address that specifically, because we have made a lot of comments on the negatives. Let's now consider this one question.

Should there be separate tables and recommendations for interpretation made for the smear positives versus the smear negatives, given that they have different statistical significance when analyzed. Let's just go around quickly.

DR. HAMMERSCHLAG: Yes.

DR. SANDERS: I am not sure I can answer that question. I am not sure how closely the package insert is read, as opposed to the technical training that goes into training laboratory personnel.

DR. CHARACHE: The responsibility is that a laboratory is able to interpret the tests they have done.

DR. WEINSTEIN: Yes.

MR. REYNOLDS: Yes.

DR. GATES: I can't answer that.

DR. RELLER: Are we responding to Dr. Gutman's question?

DR. CHARACHE: We are responding to the question of, should there be separate information provided or required for the smear positives and the smear negatives?

DR. RELLER: I think that, even though the data

base is small, that there are unequivocal differences between a positive and a negative, with regard to sensitivity.

DR. SPECTER: Yes, and primarily because there needs to be something to highlight the differences.

DR. O'BRIEN: Yes.

DR. EDELSTEIN: Yes.

DR. TUAZON: Yes.

DR. WILSON: Yes.

DR. CHARACHE: So, there is agreement that there has to be differentiation made.

Item 3, we have already discussed many of the points that come through on this one. So, we are going to limit our discussion to 15 or 20 minutes.

Does the current study, plus the data and information -- does the current study, plus data and information from previous studies, provide sufficient evidence to modify current labeling as requested by the applicant?

That means, do we have enough data, at the present time, on the smear negatives, to permit the overall conclusions that it can be used. Do we have enough data so that it could be said at this time?

DR. HAMMERSCHLAG: I must say that there is

never enough data, but I would say probably, with the qualifications that we have previously discussed. This is getting kind of redundant.

DR. SANDERS: I concur.

MR. WEINSTEIN: I think I concur. Is the statistician still here? I ask only because we talk about how difficult it is to get the data on the smear negative, culture positive.

We have said that we have got the best data that we can get for the moment. I would just like to know that the statisticians feel comfortable with that.

DR. GUTMAN: Our statistician will be back in the room shortly, so you will have to ask the question again.

DR. CHARACHE: I think we have to separate the issue of whether it is necessary to do more studies, or whether it is necessary to analyze the data for the smear negatives, so that that data can be examined.

The earlier discussion suggested that we should analyze it by study site, that we should analyze it by the definition of clinical positive diagnosis.

So, we are not recommending to say that, if we need more data, it means more clinical trials.

DR. GUTMAN: Let me remind you that, if you ask

for more data, it could be either pre-approval or post-approval. You have both options.

DR. WEINSTEIN: It is just the way the question is worded. Does the current study provide sufficient evidence to modify current labeling.

My concern, really, is just that the N is small. We all understand at this point that it is the best we can do under the circumstances.

I would just like to hear from the statistician whether or not the statistician is comfortable with those small numbers.

MR. DAWSON: Yes.

DR. WEINSTEIN: My vote is yes.

MR. REYNOLDS: Yes.

DR. GATES: Yes.

DR. RELLER: Can I ask a question?

DR. CHARACHE: Sure.

DR. RELLER: For the patients who, unfortunately, had the initial specimen with the bad primer, were those the only ones that were retested with the frozen specimens, or were there some others?

What I am really getting at is, if there had not been a glitch and everything was sort of going all the way, the number, the proportion, et cetera, would we have

twice as many patients, half again as many patients, et cetera.

Just so we have the data base and don't get hung up on what the number might have to be.

MR. MATHEWS: May I suggest that there were a few patients who were not available for retesting. There were some patients who were initially tested with the bad primer that we could not retest.

DR. CHARACHE: I am sorry, but if we don't want to be left here when the lights go out, we really have to move along.

DR. RELLER: That doesn't answer the question. I realize that -- you say there was a small number that you wanted to retest that you couldn't.

MR. MATHEWS: Correct.

DR. RELLER: Were the ones that you did retest all from patients that you wanted to retest because of the bad primer? That is what I am getting at.

MR. MATHEWS: Yes.

DR. RELLER: Basically, you could, for numbers purposes, of the 17 that we know were culture positive, and for the others, there is a high probability of smear positive, although based on an ultimate clinical diagnosis.

You get some idea for the number of patients in the earlier study that we are dealing with. Do you see what I mean?

DR. CHARACHE: I think we do have to move along. How do you vote on that?

DR. RELLER: I do not think that the numbers are adequate on the smear negatives.

DR. CHARACHE: Dr. Specter?

DR. SPECTER: Yes.

DR. O'BRIEN: As worded, no, because the applicant is asking for a smear-independent labeling indication. I think all our discussions indicated that we think smear-negative patients merit special consideration.

Along with what we have been discussing, I would then answer yes, but with that proviso.

DR. EDELSTEIN: Yes.

DR. TUAZON: Yes.

DR. WILSON: No.

DR. CHARACHE: All right, I will just ask for comments on the last no, because it was our panel leader who reviewed it. Will you comment?

DR. WILSON: Basically, what we are concerned about here is the smear negative patients, with 17 of 27

patients to look at.

Clearly, the sensitivity is only on about two-thirds of these patients, and the specificity is higher.

But I don't think that an N of 17 is sufficient evidence to modify the current labeling as requested.

DR. CHARACHE: I will again express a desire for more information about how these patients were classified as a clinical diagnosis.

They were all culture positive, and I would like to know if there were identical patients who were culture negative, where the clinical decision was only based on the culture.

I would favor looking at some more information. I am not sure whether I favor that before or after release of the drug.

Could we look at item number one under A, rather than A globally, which is to say, should MTD testing of smear-negative specimens be indicated in selected patients, those with a high clinical suspicion, or should it be open without any suggestion that that be followed.

Could you just go around quickly again, based on what we have said before? Should recommendations be made to limit the testing to selected patients, or to anyone who now gets a culture?

DR. HAMMERSCHLAG: I have a little problem with that. Again, it gets into the issue of classifying whether a patient should be tested.

I would like to see it done in those patients who are thought to be at higher risk and have a clinical suspicion, but the thing is how do you define it. My high risk may not be your high risk. So, I don't know how we can do that.

I think that if we have the previous limitations and comments, that may cover it.

DR. SANDERS: I would like to see it as a global smear in either smear positive or smear negative patients.

DR. WEINSTEIN: I actually would answer this question yes. I think that most clinicians, dealing with patients that they suspect have tuberculosis, intuitively know which ones they have the highest clinical suspicion of.

Every once in a while, we are going to be surprised, but I think that I would answer that question yes.

MR. REYNOLDS: I concur with Dr. Weinstein.

DR. GATES: No.

DR. RELLER: I think the test needs to be more

selective.

DR. SPECTER: No.

DR. O'BRIEN: Yes.

DR. EDELSTEIN: No.

DR. TUAZON: Yes.

DR. WILSON: Yes.

DR. CHARACHE: The panel is pretty divided.

Should labeling link the use of the MTD to testing smear-negative specimens in high-prevalence settings. I think we may have covered that.

DR. HAMMERSCHLAG: That has essentially been covered in previous discussions.

DR. CHARACHE: I think the discussion suggested that it should be more targeted toward individual patients and the risks explained to the clinical user.

If no, what additional data and analysis, I think we have also answered that question previously.

I think it is time to move on, with your permission to a vote.

Anyone from the public who wishes to address the group at this time?

Hearing no, we can go on to a vote. Does FDA have a response? No. Does industry have a response? Now we can move on to a vote.

MS. POOLE: The panel has three different choices or recommendations to make. The first is approval with no attached conditions.

The second choice, approvable subject to specified conditions and, prior to the vote, all those conditions should be discussed by the panel and listed by the panel chair, or third, not approvable.

The Food, Drug and Cosmetic Act specifies five reasons for denial of approval. Three apply to panel deliberations.

The reasons for voting not approvable are safety -- the data do not provide reasonable assurance that the device is safe under the conditions of use described, recommended or suggested in the proposed labeling -- effectiveness -- reasonable assurance has not been given that the device is effective under conditions of use in the labeling.

Third, labeling, based on a fair evaluation of all the material facts and your discussions, you believe the proposed labeling to be false or misleading.

The basis for safety is valid scientific evidence. It must demonstrate probable benefits for health outweighing any possible risks under conditions of use, and it must demonstrate an absence of unreasonable

risk associated with use of the product.

Effectiveness, there must be valid scientific evidence and reasonable assurance should be demonstrated, that the device is effectiveness within a significant portion of the targeted population.

The use of the device, its intended uses and conditions of a use when labeled, will provide clinically significant results.

Valid scientific evidence is described as well controlled investigations, partially controlled studies, studies and objective trials without matched controls, well-documented case histories conducted by qualified experts, and reports of significant human experience with a marketed device.

Voting today are our voting members. Our voting members are Dr. Margaret Hammerschlag, Dr. Natalie Sanders, Dr. Carmelita Tuazon, Dr. Melvin Weinstein, Dr. Michael Wilson, and as a temporary voting member, Dr. Paul Edelstein.

DR. CHARACHE: We have heard the five criteria, which inform the basis of our vote. Should we take them one at a time as a group, or should we make a motion for how you wish to proceed with the vote?

MS. POOLE: You make a motion.

DR. CHARACHE: Before we discuss. All right, may we have a motion as to approval? We need a motion whether we are voting for approval, approval with conditions or disapproval. We will ask Dr. Wilson to make a motion.

DR. WILSON: I would like to make a motion for approval with conditions, and the conditions I would like to propose are those that we have already discussed during the presentation today, in terms of the labeling.

DR. CHARACHE: Do I hear a second?

DR. HAMMERSCHLAG: I would second it.

DR. CHARACHE: Now we open it for discussion. We have to list the conditions. Perhaps one way of getting such a list is to consider these five points that were raised from the discussion. What are our conditions?

DR. SANDERS: One of the conditions was to have a prevalence table, or a warning regarding the use of this test in smear-negative patients, and the use of this test in low prevalence areas.

DR. CHARACHE: I think there was a condition that the graph itself be included along with advice on how to interpret the graph.

DR. HAMMERSCHLAG: It would also include a

warning that the data presented is based on a population prevalence of one percent, and basically is like these investment commercials, that you may not have the same results.

DR. SANDERS: Another condition was, if I heard correctly, that this test must be used in conjunction with culture.

DR. CHARACHE: Other suggestions, recommendations, remembrances of things past?

DR. EDELSTEIN: That discussion of test performance be separated for smear negative and smear positive patients, and that 95 percent confidence bands be included in the analysis of predictive value.

I would also like to add another condition which we didn't discuss, which is that post-marketing studies be conducted to determine the prevalence of the test, and that these be reported to FDA on an annual basis, and that FDA, on the basis of review of these data, may make a decision as to whether to modify or propose modification in the product labeling.

DR. CHARACHE: Would you want to also request that the full information be provided on the fresh specimens in smear-negative cases, according to the same types of analysis that we have seen in the population as

a whole?

DR. EDELSTEIN: Yes, I would.

DR. CHARACHE: Other thoughts?

If there are no additional thoughts people would like to add, we will now go around the panel, for the voting members, ask that you first vote, and then explain the reasons for your vote for the record.

DR. HAMMERSCHLAG: Okay, I vote to accept the motion for approval with conditions. I concur with the conditions, for the reasons that we discussed.

I think that the test has value if we understand its limitations. That has to be communicated to the people who will use it.

DR. SANDERS: I also vote for approval with conditions, for the reasons said already.

DR. WEINSTEIN: I, too, vote in favor with those conditions, for the same reason as the previous two voters.

DR. EDELSTEIN: I vote for approval with conditions as stipulated, for the reasons that interpretation of the test can be made, given the data about pretest probabilities.

DR. TUAZON: I vote for approval of the device with conditions listed.

DR. WILSON: I vote for approval as stated.

DR. CHARACHE: Thank you.

MS. POOLE: Thank you very much for your time and your efforts, and if there are no more questions, don't forget, take everything that you brought with you. The panel, you can leave your submissions here and we will take care of them.

DR. CHARACHE: I would also like to thank everybody for their very thoughtful and intensely concentrated efforts. It has been a long day. Thank you.

[Whereupon, at 7:02 p.m., the meeting was adjourned.]