

Medical Devices Advisory Committee
Federal Drug Administration

Open Meeting
of the
Hematology and Pathology
Devices Panel

Wednesday, January 28, 1998

Conference Rooms G & H
5600 Fishers Lane
Rockville, Maryland 20857

Proceedings By:

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

PARTICIPANTS LIST

Acting Chairperson

John L. Francis, Ph.D.

Voting Members

George G. Birdsong, M.D.
Diane D. Davey, M.D.

Temporary Voting Members

Machelle Harris-Allen, M.D.
Juan C. Felix, M.D.
Blake R. Nestok, M.D.
Andrew A. Renshaw, M.D.
Shailini Singh, M.D.
Robert L. Williams, Jr., M.D.

Industry Rep

Alton D. Floyd, Ph.D.

Consumer Rep

Ellen S. Rosenthal, M.S.

Nonvoting Consultant

Ezra C. Davidson, Jr., M.D.

PRESENTERS

David M. Garner, Ph.D.
Sandra K. Fite, CT(ASCP)
Alan M. Kaye
Joan L. Shaver, Ph.D., R.N., F.A.A.N.
Carol Ann Armenti
Paul T. Wertlake, M.D.
Michael Stanley, M.D.(ASCP)
Kara L. Anderson, RNC, B.A.
Diana Samia Silman

Written Statements

American Social Health Association (ASHA)
The Trylon Corporation
Society for the Advancement of Women's Health Research
(SAWHR)

Table of Contents

Opening Remarks -- Introductions	1
Ms. Calvin	
Open Public Hearing	
Dr. Francis	6
Sponsor Presentation - NeoPath, Inc.	
Ms. Norton	56
Dr. Nelson	63
Dr. Wilbur	69
Dr. Bonfiglio	85
FDA Presentation	
Ms. Stuart	97
Ms. Chen	101
Open Committee Discussion	
Dr. Francis	104
Panel Vote and Recommendations to FDA	
Dr. Francis	162
Adjournment	
Dr. Francis	193

P R O C E E D I N G S

(10:15 a.m.)

Agenda Item: Opening Remarks -- Introduction

MS. CALVIN: Good morning and welcome to the Hematology and Pathology Devices Panel meeting. I am Veronica Calvin, Executive Secretary for the Panel.

Before we begin today's agenda, I will read the summary minutes from the last meeting. The last meeting of the Hematology and Pathology Devices Panel was held on Friday, September 5, 1997. The committee discussed quality control issues for home-use prothrombin time devices.

The FDA requested input on appropriate QC procedures and recommendations for data requirements. A number of useful comments were received from the Panel members, guest speaker, Dr. Barbara Gail Macik, public speakers, and industry, and will be considered in the review of future prothrombin time device submissions.

Today the Committee will discuss a premarket approval, or PMA, supplement for a computerized automated Pap smear reader that is indicated for use as a primary screener to select a subpopulation of smears that will be designated for no further review. Attached to your agenda you should find the specific questions to be addressed during the Open Committee Discussion, which will take place in the afternoon session.

At this time I would like to introduce our Acting Panel Chairman, Dr. John Francis. Dr. Francis is the Director of Cancer Research at Walt Disney Memorial Cancer Institute in Florida. I would also like to introduce Dr. Steven Gutman, who is the Director of the Division of Clinical Laboratory Devices in the Office of Device Evaluation.

Now I would like for the panel members to introduce themselves, beginning with Dr. Floyd. Please state your name and affiliation.

DR. FLOYD: Alton D. Floyd, Ph.D. I am President of Trigon Technology and the industry representative on this Panel.

MS. ROSENTHAL: Ellen Rosenthal. I am the consumer representative to this Panel. I am an engineer and a freelance writer.

DR. FELIX: My name is Juan C. Felix. I am member of the Advisory Panel, USC.

DR. RENSHAW: I am Andrew Renshaw. I am member of the Advisory Panel. I am from Brigham and Women's Hospital in Boston and Harvard Medical School.

DR. WILLIAMS: I am Robert L. Williams, Jr. I am a physician in Atlanta, Georgia. I am the Chairman of the OB Department at Southwest Hospital, and I am on the

Advisory Panel.

DR. DAVEY: Diane Davey. I am a Panel member. I am Director of Cytopathology, University of Kentucky, in Lexington.

DR. NESTOK: Dr. Blake Nestok, Medical Director of Cytology, Christ Hospital, Cincinnati, and Health Alliance, Cincinnati.

DR. BIRDSONG: George Birdsong, M.D. I am a Panel member, Director of Cytology at Grady Memorial Hospital in Atlanta, Georgia, and Pathology faculty at Emory University in Atlanta.

DR. DAVIDSON: Ezra Davidson. I am Professor of Obstetrics and Gynecology at the Drew University of Medicine and Science at UCLA, and I am consultant to the Panel today.

DR. GUTMAN: I am Steve Gutman. I am the Director of the Division.

MS. CALVIN: Thank you. For the record read the Conflict of Interest Statement.

The following announcement addresses conflict of interest issues associated with this meeting and is made part of the record to preclude even the appearance of a 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 employers' 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 in the best interest of the government.

We would like to note for the record that Dr. John Francis has consented to serve as chairman in the absence of Dr. Timothy O'Leary.

We would also like to note for the record that the Agency took into consideration certain matters regarding Drs. Diane Davey and George Birdsong. Dr. Birdsong reported contracts and speaking engagements with firms at issue. Since the contracts are not related to today's agenda and the educational lectures addressed general issues for which he received no compensation, the Agency has determined that he may participate in today's deliberations. Dr. Davey reported a past involvement with firms at issue; however, since these involvements did not involve Dr. Davey directly, the Agency has determined that Dr. Davey may participate in the Committee's deliberations.

In the event that the discussions involve any other products or firms not already on the agenda for which an FDA participant has a financial interest, the participants should excuse themselves from such involvement and their 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.

I will now read the Appointment to Temporary Voting Status.

Pursuant to the authority granted under the Medical Devices Advisory Committee Charter, dated October 27, 1990, I appoint the following people as voting members of the Subcommittee of the Hematology and Pathology Devices Panel for the duration of this Panel meeting on January 28, 1998: Mabelle H. Allen, M.D., Juan C. Felix, M.D., Blake R. Nestok, M.D., Andrew A. Renshaw, M.D., Shailini Singh, M.D., Robert L. Williams, Jr., M.D.

For the record, these people are special government employees and are either a consultant to this Panel, or a consultant or voting member of another panel under the Medical Devices Advisory Committee. They have

undergone the customary conflict of interest review. They have reviewed the material to be considered at this meeting.

Signed, D. Bruce Burlington, M.D., Director,
Center for Devices and Radiological Health. Dated January
15, 1998.

Now, I will turn the meeting over to Dr. Francis.

DR. FRANCIS: Thank you, Ms. Calvin. I would like now to start the open public hearing. We have a very full agenda. There are nine individuals who have indicated they wish to address this meeting. They have indicated while making that request what length of time they will need, and I plan to hold them very tightly to that indication. And I would remind each presenter that before they address the Panel, state whether or not they have any financial involvement with the manufacturers of any products being discussed, or with the competitors.

With that in mind, I would like to start the proceedings by inviting Dr. David Garner, Senior Scientist, British Columbia Cancer Agency, Vancouver, Canada, to step up to the podium.

Agenda Item: Open Public Hearing

DR. GARNER: Good morning, ladies and gentlemen. My name is David Garner. I am a senior scientist at the British Columbia Cancer Research Center in Vancouver, and I

must say it is nice to have a chance to escape the rain of Vancouver to be here.

I am supposed to disclose, my expenses for this trip will be paid for by the British Columbia Cancer Agency. One of my duties with the Agency is to assist in the assessment of new technologies, particularly related to screening, not just of cervical cancer but for all cancers.

I should disclose that between 1990 and late 1996, I worked for a company called Zillex(?) Technologies, doing among other things, computer-assisted Pap screening.

What I want to talk to you about today is a very small point, but I think it is a relative point, about Pap prescreeners. I think that some people have the notion that a Pap prescreener can have any level of accuracy, that it may increase both the sensitivity and specificity.

It turns out that the underlying logic of prescreening does set limits, and those limits can be calculated. I do not have time to go through the proof of those calculations, but I have written a paper which I have submitted to Ms. Calvin, and I have extra copies if anyone needs them. And so, what I am going to focus on is what the answer is, and what the implications are.

Quickly, let's look at what a prescreener is. I apologize for this kind of busy diagram. The circles, the

colored circles -- represent cases, the green ones, of course, being negative, and the red ones being positive. And Test A is performed.

In this case, it is a device which sorts them into two groups, the right, green ones, being negative, and you can see there are some false negative errors. And the left, red ones being positive. The right ones being labeled, no further review indicated, and the left ones being, need further screening.

The second test is performed by the cytotechnologist, which again sorts the slides, and so the total negatives are given by the sum of these.

The key point here, actually, is that in order for a slide to be called positive, it must be called positive both by the machine -- by Test A -- and by the cytotechnologist. It is a logical and.

It turns out when you -- don't worry, I am not going to go through this whole box of slides -- when you work through this whole thing, this is the result that comes out, and in fact, I have also shown it for a post-screener. This is the quality control application, which has already been approved by FDA several years ago.

The point I want to draw your attention to is that, the sensitivity of a prescreener is bounded by these

two values, and the real point is that the maximum sensitivity that it can have is equal to the sensitivity of the least sensitive of the two-component test.

On the other hand, specificity can improve tremendously; it can go up to 100 percent, and the specificity is at least as good as that of the most specific test.

If you take a look at the case of a post-screener, you will see that there are some quite pretty symmetry here, and you get the opposite tradeoff. You can improve sensitivity, at the cost of specificity. But, the point I want to draw here is that, for a Pap prescreener, the sensitivity is always less than or equal to the least sensitive of Tests A and B; the specificity is always greater than, or equal to, the most specific.

Now, not to put too fine of a point on it, I do not want to leave you with the impression that the sensitivity will collapse; not necessarily true, and in fact it is unlikely. So, here are -- I just plugged in some numbers. This is where a machine -- I do not know whether this is what the manufacturer before you today will show, but it is typically something like this; in excess of 90 percent sensitivity, at some kind of a so-called split ratio.

Perhaps cytotechs are 85 percent sensitive and their specificity is probably somewhat controversial -- I think it is 90 percent. But anyway, if you plug the numbers in, you will see that the sensitivity can increase -- it is cannot be higher than the 85 percent of the lowest, but it is not much lower. So, I do not want to put too fine of a point on it, it is not like sensitivity will collapse.

It turns out that the underlying logic of a Pap prescreener can be found in other completely unrelated tests, and I just want to show an example here. This is from New England Journal of Medicine about a year ago on a fecal occult blood screen for colon cancer. And it did exactly the same thing.

All I want to draw your attention to is the fact that the numbers come out the way I predicted; that is, the sensitivity of the combined test is less than the sensitivity of the worst test. And conversely, the specificity of the combined test is better than the specificity of the best individual test. Here is an example for bladder cancer doing the OR.

One other result that comes out of this, which I will not go through, really, is that the tests are commutative. It does not matter what order the tests are performed in, the accuracy is independent of that. That is

not intuitively obvious, but it has great implications for cost-effectiveness. Cost-effectiveness, of course, very much depends on the order of the tests.

Well, all calculations depend on assumptions, and there are assumptions and this, and so what are the assumptions? The key assumption is that the two tests are independent of each other in the sense that one test does not affect the other.

Now, the most easy way for one test to affect the other is by ruining the sample, and that is not the case with this manufacturer, as I understand it; these are Pap smears and they are being looked at twice. But there is the possibility of this issue of vigilance. It is unlikely that the machine is influenced by the people, but the people could be influenced by the machine. And so, what is the result?

Well, there are three possibilities. The people could stay the same. They could get worse, or they could get better. If they stay the same, then what I said is true, the calculations I have already presented.

If they get worse, then what I said is optimistic, and sensitivity certainly is not going to improve.

If the cytotechnologists improve, then the sensitivity of the overall test could improve, so long as

they improve by at least as much as the false negative proportion of the machine. So, again, just to quickly remind you of this diagram I showed you earlier. The cytotechs have to somehow make up for these mistakes.

Okay, to draw this to a close. Conclusions. In principle, a Pap prescreener can greatly improve the Pap test's specificity. Usually, that will result in improvement in the positive predictive value, and that may lead to a great improvement in cost-effectiveness. However, a Pap prescreener cannot improve the sensitivity of the Pap test, except under these circumstances of changed vigilance.

Changed vigilance does present its problems, because it is probably something which is both individual-specific, and time-dependent.

Thank you very much for your attention.

DR. FRANCIS: Thank you, Dr. Garner for your presentation and keeping to time. In the interest of time, I am going to have all the presentations first and if there is time after those, and after reading the statements submitted, then I will call for comments and discussions from the Panel and anybody else.

I would now like to invite the next speaker, Sandra K. Fite, Cytopathology Supervisor, Chapell-Joyce Pathology Association, Texarkana, Texas. Step up to the

podium, please. And I would remind you to state whether or not you have any financial involvement with the manufacturers of any products being discussed or with their competitors. Thank you.

MS. FITE: Distinguished Chairman, and members of the Advisory Panel, my name is Sandra K. Fite, CT(ASCP)I am a Cytology Supervisor at Chappell-Joyce Pathology, Texarkana, Texas. NeoPath is paying my expenses for this trip, however, I am here on my own time, and I am also a NeoPath stockholder.

Over 30 years ago, just after completing cytology school, I started hearing that I was going to be replaced by a machine that could screen Pap smears. About eight months ago, when I saw the results of the clinical trials of the AutoPap QC System, I knew the time had finally come.

At first, I was very impressed with the results, and I was excited about the new technology knew that the primary system would be even better than the QC system at reducing the false negative rate. Then reality set in. I really could be replaced by a machine.

Cytology is on the verge of becoming fully-automated, and the changes that such automation bring to the cytology community will be felt throughout. Some cytotechnologists -- such as myself -- are preparing to work

alongside automation; some are vigorously fighting the new technology; and others are just fearful of losing their jobs.

Recently, a cytotechnologist at the laboratory where I work asked me if I thought she was going to be replaced by a machine. I said, maybe, but not for many years to come. Cytotechnologists with her skills will be needed to meticulously screen and diagnose those Pap smears that the AutoPap System selects for manual review.

Automation will not replace good cytotechnologists, it will make us more accurate Pap smear screeners.

Today, there is a shortage of good, experienced cytotechnologists. Many stay in the profession for only a few years. When I was in cytology school, there were five in my class. Within ten years, there were only two left in the field of cytology, and within 15 years, I was the only one left.

Certain urban geographic areas are particularly short of qualified cytotechnologists. My laboratory just hired a third one after an eight-month search. Finding a skilled cytotechnologist that is willing to relocate is extremely difficult. Because we are not in a metropolitan area, we cannot easily attract cytotechnologists; however,

many metropolitan areas are also faced with a shortage. Recently, a large national laboratory had a backlog of over 10,000 Pap smears. To catch up, the cytotechnologists were required to work seven days a week. Screening day after day, week after week, without a day off is not an optimal situation.

Under such conditions, mistakes are more apt to occur. The AutoPap System can help relieve the shortage of cytotechnologists by selecting a portion that do not require manual review. Now, true, in certain geographic areas of the United States, there is not a shortage of cytotechnologists, and there are some cytotechnologists looking for jobs.

One reason for the shortage may be that many leave the profession because of boredom. The day-to-day job of looking through a microscope, screening primarily negative Pap smears, is an extremely repetitive, boring job. This boredom may also increase the false negative rate.

The AutoPap System in our laboratories can help eliminate the boredom and repetition associated with Pap smear screening by leaving the more challenging ones for cytotechnologists to screen.

I have been a cytotechnologist for over 33 years. For most of my career, I have worked as a cytology

supervisor and an anatomical pathology laboratory manager in large Dallas laboratories. Currently, I am the Cytology Supervisor in a northeast Texas laboratory.

During my career, I have screen hundreds of thousands of Pap smears, and rescreened tens of thousands initially screened by other cytotechnologists. This rescreening was done for quality control and to assess the screening performance of other cytotechnologists within the laboratory.

From this experience, I know that not all cytotechnologists have the same ability to detect abnormal cells, and even the very best miss abnormal cells from time to time.

Most of us are doing an excellent job, but even the very best of us make mistakes. Three weeks ago, the AutoPap QC System was delivered to my laboratory. During a recent e-mail discussion with a well-known cytopathologist, I was asked, how much it is going to cost my laboratory to find a false negative with the AutoPap System?

My answer was, I really do not know. And then later, when I got to thinking about it, I do not think I could put a price tag on a woman's life. A missed Pap smear can be devastating to the patient. It can cost a woman her life. And this is what the primary screener is all about;

saving women's lives.

Some think the time has not yet come; I think it has been too long coming. We need the AutoPap Primary System in our in our laboratories now. Thank you.

DR. FRANCIS: Thank you, Ms. Fite. Now, I would like to ask Alan Kaye, Executive Director of the National Cervical Cancer Coalition, PathNet Esoteric Laboratory Institute, Van Nuys, California, to address the Panel. Once again, I would remind the speaker to disclose any financial involvement with any of the manufactures at issue or with their competitors.

MR. KAYE: I have no financial involvement with any of the companies or any of the competitors. Before me is a document I received last week to read about a woman, whose name is Wendy, who is dying of cervical cancer.

She had seven Pap smears in a row, one year apart, yet she is dying of cervical cancer. It is possible that the Pap smears she had may have missed the cancer.

I am before you today as the Executive Director of the National Cervical Cancer Coalition. Our coalition is made up of women and groups, fellow citizens, pathologists, cytotechnologists, laboratorians throughout the country.

I am also the President and CEO of a national cytopathology practice laboratory. As a person closely

involved in laboratory medicine for over 25 years, I am able to bring some practical marketplace experiences into my position as Executive Director of the National Cervical Cancer Coalition.

Our coalition is here today to urge you -- the FDA Medical Advisory Panel -- that if the scientific data before you is significant, and statistically valid, that you must recommend this new technology to the FDA for immediate approval.

As you know, the accuracy and reliability of cervical cancer screening is a major public health concern. As the most important tool in the prevention of cervical cancer, an accurate Pap smear is vital to promoting women's health.

Increasing access to regular Pap smear screening has long been a matter of reimbursement economics and public health policy.

Our coalition wants to make sure the new cervical cancer screening technologies that are scientifically sound, will be approved, reimbursed properly, and available to all women, especially to the women that are most in need of the Pap smear.

New technology has moved into laboratories for a long time now. I can remember in hematology, when the CBC,

the complete blood count, had a manual differential slide, and was looked at by medical technologists in the laboratory.

First, a slide had to have a light feather smear of blood spread over the glass slide, just the right way. Then the slide would come into the laboratory for processing. A department full of medical technologists sitting at their microscopes with the cell counters would process each slide individually, counting the blood cells and then multiplying out the number to approximately determine the amount of blood cells seen.

Automation through robotics and information technology have been able to automate the differential portion of the CBC, and for that matter, the entire CBC. We now have hematology analyzers that completely automate the most time-consuming tasks, improve accuracy over old manual methods, and allow for more efficient pricing.

When run properly, these hematology analyzers can assume liability for the accuracy of the results they produce, and after a period of about ten days, I believe the CLIA guidelines allow for the specimens to be discarded.

I believe the CLIA guidelines will probably change when it comes to cervical cancer testing, also. New technology is now needed to move into the area of cervical

cancer testing.

I am not -- and I repeat -- I am not suggesting for a minute that the Pap smear will ever be as automated as the CBC. The Pap smear is classified as a high complexity procedure. This means that inaccurate or improperly interpreted tests could result in significant harm to the patient. Employing qualified personnel to perform the tests, and adhering to approved, quality control procedures, is crucial; however, if the current science before you indicates that there is a method available that can automate and sort out a portion of the Pap smears that are within normal limits, you must consider this technology.

By approving this technology, you will be able to allow the qualified laboratorians to focus in on the 75 percent of the cases that may always require a human interpretation and judgment. Imagine, 25 percent of the cases may be classified within normal limits.

These cases, just like the CBC specimens, may have liability for accuracy stop with the equipment manufacture. It should be stated in a package insert. Also, on these same specimens, the accuracy is so high that CLIA will need to immediately modify its regulation stating that any slides classified within normal limits for this new technology, need only be maintained for ten days to two

months or so after the report is issued. This will reduce costs to the laboratory for slide storage and handling, and allow for improved pricing for women to have greater access to the Pap smear.

We are all aware of the various new technologies that are now surfacing for improving cervical cancer screening and eliminating the disease. There are some excellent thin layer specimen preparation methods, HPV detection probes, and off on the horizon, HPV vaccine.

It is very possible that in the near future, we may see a combination of thin layer specimen preparation, and computer-assisted technology coming together; however, this possibility will not occur if you do not recommend to the FDA for approval of the technology before you today.

We need to get this technology into the marketplace, allow women and their physicians to decide what technology is best for their needs.

This technology provides very useful data for enhancing the laboratory's ability to screen Pap smears. In the Information Age in which we live, it would be ludicrous to not allow this technology to go forward into the marketplace. After all, this technology will help provide the labs with additional diagnostic data for their Pap smears. Data, information, and knowledge. The data becomes

the information we need to develop the knowledge for better population-based risk management of cervical cancer, and most importantly, for improving patient outcomes. That is true managed care, improving the patient outcomes.

You have a unique opportunity before you to improve patient outcomes and help begin to further reduce false negative rates of cervical cancer in women. Nobody wants to read more stories about women like Wendy that I first talked about, who have gone for their regular Pap smears, and are needlessly dying of cervical cancer.

On behalf of the National Cervical Cancer Coalition, I am hopeful and confident you will recommend to the FDA to allow this technology to immediately move forward for FDA approval.

You are in a position to help reduce the cervical cancer rate among women in this country, beginning now. God bless you on the recommendation you have before you to make.

DR. FRANCIS: Thank you, Mr. Kaye. I would now like to ask Dr. Joan L. Shaver, Dean of College of Nursing, University of Illinois at Chicago to step up to the podium and address the Committee. Please state whether you have any financial involvement with the manufacturer of products being discussed or with their competitors.

DR. SHAVER: Thank you very much for the

opportunity to be here to present my views, which are my own views, and I have no financial interest in any of the companies you mentioned.

Women's health is my professional interest. My name is Joan Shaver. I am a researcher of women's health and I am currently a Professor and Dean of the College of Nursing at the University of Illinois at Chicago. I am affiliated with the UIC Center on Research on Women and Gender.

As a midlife woman, I am a consumer, of course, of this technology. Furthermore, I am a nurse, and nursing is a female-dominated profession, so as nurses, we have a particular interest in women's health issues as it affects the health of our female patients, most of us, personally, and as colleagues.

I urge serious consideration of this technology to increase the capacity overall for better women's health services. The importance of cervical cancer is clear as the seventh most common cause of death from cancer in the United States, but the most common cause of cancer in women around the world, as a matter of fact.

The Pap smear is an extraordinarily effective means to effect a high cure rate by early diagnosis, but only if it is available on a wide scale basis to those most

in need.

I am concerned that we strive to be thinking of this test as it was originally envisioned; that is, a screening test for the population, rather than a diagnostic test for individuals, and that we think about regulating the technology applicable to this test from the public health perspective.

The tradeoffs I think that we might make between our considerations of quality, access, and cost will differ, depending on those perspectives.

Litigation is making it economically difficult to serve the women who need this test, which in fact, is all women. The Pap smear as a public health screening device versus a diagnostic device having to meet highly-regulated and rigorous standards for which costs threaten to make it available only to the financially advantaged, is not in the best interest of women in general.

It has been said that it is more cost-effective to increase the frequency of testing and keep the price low, than to decrease false negative rates and increase the price. And in fact, from a public health perspective, the capacity to do this test, therefore, presumably the cost to do this test, dictates access.

A public health screening technology must be

without question safe and as accurate as we can make it. So, the high standards of performance compared to the human scoring are a necessity for this technology being considered.

The data that I have seen vis-a-vis the NeoPath technology points in the direction of being very promising with respect to the accuracy or quality of the test. But the issues are also that a public health screening device must be rapid and relatively simple.

Some arguments have already been made about the relative simplicity tasks, when one does large samples are repetitive and tedious, and this provokes errors in terms of human scoring. Using humans to complete all parts of this kind of a testing adds complexities that the automation can in fact help us alleviate.

It has also been mentioned there is inconsistency of performance over time, dependent on human motivation. Inter-individual competencies, in general, are an issue; boredom and fatigue are an issue. This has led, in the reading I have done, for us to institute various complex strategies for quality control, which have the, I think, potential to be very costly.

These include: retrospective screening; random rescreening; targeted rescreening; sequential block

rescreening; workload adjustments; environmental modifications; error calculations and personal feedback; overscoring; and so on.

A strictly human process for a public health screening element spawns the need for training, retraining, continuing education, and there is the potential to reduce some of this complexity using automation of this process. That means we should be able to do it at a higher rate around the clock and relatively less supervised.

If safety is not an issue, at least in the testing part, and if when used under realistic conditions, the technology performs tasks of comparable quality, then there should be cost advantages to be achieved through increased capacity with less overhead.

It can be argued that there is potential in this technology for increasing capacity, and actually, the world looks to us I think for guidance in many of our public health endeavors to improve health, and we have the potential, in fact, to provide something to the rest of the world, where in fact human training to do this kind of screening is not anywhere near as great a potential as here, or as good quality.

Increased cost-effective capacity means screening access can be stepped up for earlier detection in more

women, leading presumably to reduced invasive -- and therefore costly -- treatments.

This means that women's health care can be better served, and I urge us to consider the technology in the perspective of improving the capacity for women's health services in general.

Thank you.

DR. FRANCIS: Thank you, Dr. Shaver for your comments. I would now like to ask Carol Ann Armenti, the Director, Center for Cervical Health, Toma River, New Jersey to address the Panel. I would like to remind you to state whether you have any financial involvement with the manufacturers of any products being discussed, or their competitors.

MS. ARMENTI: I have no financial interest in NeoPath or in the outcome of these proceedings. My expenses were paid by NeoPath so that I could testify to --

DR. FRANCIS: Could you just pull the microphone over to you?

MS. ARMENTI: Yes, certainly. Good morning, Chairman, and members of the Panel. My name is Carol Ann Armenti and I survived cervical cancer.

I come here today to speak for myself and as Director of the Center for Cervical Health. Perhaps more

importantly, I come here to address you on behalf of those women who are not with us today -- Dolores Gary, late of Wisconsin; Rochelle Safieri, late of Pennsylvania; and the 5,000 women who did not survive cervical cancer, in this country, last year.

I speak for Tina, who faces an uncertain future in Georgia, and for Laura in Washington, whose most ardent hope is to live to see her daughter's fourth birthday. And the 10,000 women who will endure and survive cervical cancer in this country this year.

I also speak for the women from all parts of this country who have called me, frightened but courageous, who have been diagnosed with cervical lesions this past year and who worry for their future.

I began receiving regular Pap tests during my college years. What those early tests indicated will never be known because my physician became suddenly ill, he discontinued his practice, and my records disappeared with him.

About that time, the Clinical Laboratory Improvement Act of 1988 was passed, which all hoped, would increase the accountability for the proper handling of Pap tests, and ensure a woman's access to her own slides.

It was hoped that physicians would provide

laboratories with better samples, and that laboratories would process these samples with a higher degree of reliability.

Nearly all of my Pap slides prior to 1993 -- the year in which I was diagnosed with cervical cancer -- were destroyed; however, the surviving reports indicate that the samples were normal, but limited by the absence of endocervical cells. I submit to you that any Pap test which lacks endocervical cells, is no Pap test at all.

The 1992 Pap test slide which was located was reported as normal, but limited, this time by blood. The submission report from my physician which accompanied the slide, indicates it was taken during my menstrual cycle. My age, a risk factor in cervical disease, was incorrect on that report. My address and telephone number were likewise incorrect. There was no unique identifier, as required by CLIA, other than the combination of my name and the name of my physician.

The same physician, however, treated other members of my family, many of which have the same name. Upon rereading in 1994, the slide clearly showed high grade lesions, or carcinoma in situ. This slide was egregiously misread.

During the course of depositions at the litigation

I instituted for medical malpractice, the laboratory's chief pathologist testified that my slide was not expeditiously processed because, quote, there was a backlog; we were shorthanded.

At the trial, the cytotechnologist who had misread my slide was not available because he suffered from malaria; we will never know his condition when he read my slide.

At the malpractice trial my physician testified that he routinely took Pap smears during menstruation because, quote, it is convenient for the woman. I assure all of you that living with cervical cancer has not been convenient for this woman.

My cervical cancer was discovered after I had bled for more than a year following an automobile accident. I had reported the bleeding to several physicians, and insisted that my gynecologist examine me on several occasions, including after a miscarriage.

In November of 1993, I hemorrhaged in the shower alone in my home. I have no recollection of dressing myself and driving myself to a physician. I do, however, recall the horror of watching my blood splash on the floor and off the walls of the shower, and the gray jelly, later identified as blood clots were swirled amid the blood at the drain.

I recall the eyes of the physician, who would not meet mine as he said, you have cervical cancer. I also recall what it was like to wake in the hospital, vomiting and in pain, to be told that my husband and I would never have the baby for which we hoped. I remember vividly the face of the oncologist who told me, you will die.

The National Institute of Health estimates that 20 percent of all Pap smears are in error; half by sample error, half by laboratory error. A decade after the promise of CLIA, one out of every five slides will be read incorrectly this year; that is, approximately ten million Pap reports, upon which women rely to plan their medical future, will be wrong. One of 20 of those women will have undiagnosed, progressive, cervical disease.

It is shameful, ten years after the enactment of CLIA that the Head of Obstetrics and Gynecology at a major hospital endorses Pap tests during menstruation. And it is equally shameful that laboratories cannot be relied upon to return inadequate tests to physicians. It is horrific that 150,000 women since the promise of CLIA will be diagnosed in this country with a completely preventable disease.

We frequently hear that half of the women stricken with cervical cancer have never had a Pap test, thereby placing the blame for the disease on the victim, rather than

where it belongs, on the educational system and medical system, which has failed to inform women of the need for testing, and which has failed to provide access to medical care.

We rarely hear the same statistic read with the emphasis upon those of us who had Pap tests, who replied upon our physicians, our laboratories, and the government, to see that our tests were accurate.

We have been failed by those from whom we sought help and protection and in whose hands we placed our lives. We repeatedly hear self-congratulatory praise for the medical community that we have decreased the death rate from cervical cancer in this country by 70 percent. This is scant comfort to those who today cling to a life of prolonged agony, for death from cervical cancer may take as long as ten years.

I carry with me a picture of my abdomen, an abdomen which has three times survived the violation of surgery; the abdomen of a body which has endured so many tests that there are no remaining veins available in my arms, or hands, for yet another test.

I sit here having refused yet another surgery recommended by my physicians, because at this moment I simply cannot contemplate any more suffering.

This country may not applaud itself for the survival of two-thirds of the women stricken with cervical cancer; mere survival should not be the goal for a disease which is preventable.

We are told that to improve the test is not cost-effectiveness. We are held hostage to a system that decrees that if we demand a more accurate test, we will jeopardize the availability of the test for all women.

It has cost me \$100,000 in medical treatment to survive. I can no longer work. I lost my business, and I nearly lost my home to foreclosure. I will require medical care for the remainder of my life. It has been cynically said, that at a medical cost of \$300,000, it is far cheaper to die than to survive.

One wonders at the financial acumen which calculated it was more cost-effective to allow me to contract cervical cancer and survive, than to improve the reliability of the test.

It is my hope that my appearance today will engender a change in attitude in this country, which will result in better health care for women. And it is with that same spirit that I began the Center for Cervical Health.

The Center was founded when I could find no organization which dealt solely with cervical issues. As a

psychologist, I felt perhaps more keenly than most the lack of emotional support for women who are suffering from a disease which society has attributed to women -- in the words of the New York Times -- who practice high risk sex.

How shameful to brand women who are already suffering by implying they are somehow to blame for their illness through morally and socially unacceptable behavior. How better to isolate these women, for who among us wishes to profess to her husband, her mother, and her children, that she is suffering because she engaged in high risk sex? It is far easier to silence us by shame than to find a solution to our problems. It is indeed more cost-effective to ostracize us than provide us with better education and more reliable testing.

Today, you have an opportunity to address the issue of better reading of Pap smears by approving automated primary reading of these same slides, as you have approved automated rescreening of slides in the past.

After the bitterness and cynicism of my experience with cervical cancer, I have found renewed hope that all women may receive a higher standard; the same standard of care through automation.

Computers know no social class or rate of reimbursement. They are not intimidated by pressure from

hospitals, physicians, and laboratories. They do not consider lobbying efforts or contract awards. And they are not subject to backlog during the holidays, nor does their attention flag at the end of a shift.

I can tell this Panel today, that my slides, which were misread by the laboratory, were properly read by both automated rescreeners previously approved by this Panel. I ask this Panel to recommend the approval of AutoPap for primary screening, as a step in the fulfillment of the intent of CLIA. Thank you.

DR. FRANCIS: Thank you, Ms. Armenti. The Panel appreciates you making the journey to present your comments.

I would now like to call upon Dr. Paul Wertlake, Chief Medical Officer of Unilab, Tarzana, California to address the Panel. Please state whether you have any financial involvement with the manufactures of any products being discussed or with their competitors.

DR. WERTLAKE: My financial involvement in these matters is limited to NeoPath providing travel expenses.

I am the Chief Medical Officer and a pathologist for Unilab, a California laboratory performing approximately one million Pap smears per year. About one-half of those Pap smears are performed in our Southern California Lab, for which I am the Medical Director.

Our laboratories are accredited by the College of American Pathologists and our Southern California Laboratory is additionally accredited by the American Society of Cytopathology.

Our Southern California Laboratory has operated with a prevailing false negative ratio of 4 percent, and we have operated with a 37 percent cytotechnologist rescreening average, historically.

We decided that we would attempt to improve our quality by reducing false negatives using the AutoPap System, and attempt to be cost-neutral by reducing our high cytotech rescreening rate. We began use of the AutoPap in February of 1997, and we are just completing our first full year.

I would like to show a slide -- [technical difficulties ensued] -- let me proceed and I will describe data to you. We compared in this experience, in a recent review, the first six months of our experience with the AutoPap System, and we compared the experience of 1997 with the targeted reviews for cytotechnologists as picked by the AutoPap System of slides more likely to be false negative slides, in comparison to what was a strategy in 1996 of a random QC.

In 1996, the number of negative slides involved in

both cases were reasonably similar, being a bit above and a bit below 200,000. The number of QC rescreened slides in 1996 was approximately 41,000, and this was reduced in 1997, using the AutoPap System, to 29,000. So, we decreased rescreens slides from 41,000 to 29,000. And in so doing, and using the targeted QC, we did increase the false negative pick-up significantly. And most critically, in my view, in terms of the high grades, we identified 29 high grades in 1997 with the AutoPap, as compared to four in six months by random QC.

That four represented rescreening of about 11,000 slides. That is a volume that a cytotech in our labs, starting on beginning of January, would occupy them to the middle of September to find a single high grade. The adjusted effectiveness on the critically high grade basis, was eleven-fold.

Now, in my view, this is a partial benefit of the case finding power of the AutoPap System, because we were operating at a 10 percent sort rate. And if we were operating under primary screening, in which I view that we would be tapping the full capability of this technology, I believe we would see better data than that. We have been doing biopsy confirmation of cases, and I would be happy to share that. I believe that this power to case find is the

single strongest indicator for approving primary screening.

A second reason is that it does allow redirection of a limited resource, cytotechnologists, to new casework, which can be women that are not presently included in our Pap screening program.

Thank you.

DR. FRANCIS: Thank you, Dr. Wertlake. I apologize for the equipment failure so that you could not show your slides.

I would now like to ask Michael Stanley, Dr. Michael Stanley, Chief of Pathology at Hennipin County Medical Center, Minneapolis, to address the Panel. Is Dr. Stanley here?

PARTICIPANT: [Comment away from microphone.]

DR. FRANCIS: I will proceed to the next presentation. I would like to ask Kara Anderson, Director of Medical Affairs, Planned Parenthood Federation of America to address the Panel.

MS. ANDERSON: Good morning. I am Kara Anderson, a nurse-practitioner in women's health care and Director of Medical Affairs for Planned Parenthood Federation of America's national office in New York City.

Planned Parenthood's 141 affiliates with more than 900 clinic sites in the United States --

DR. FRANCIS: Dr. Anderson, I am sorry to interrupt you. Could you please state whether you have any financial interests before you start?

MS. ANDERSON: I do not.

DR. FRANCIS: Thank you.

MS. ANDERSON: Sorry. Planned Parenthood's 141 affiliates with more than 900 clinic sites in the United States do more than 1.5 million Pap smears each year.

We are constantly aware of the need to maintain quality screening at an affordable cost. False readings, positive or negative, are at best expensive and time-consuming to follow-up, and at worst, can be catastrophic.

As a woman's health care nurse-practitioner, I have cared for women like Carol Ann Armenti. I also, 25 years ago, had a series of Pap smears that were read as, quote, inflammation, unquote, but otherwise negative. Fortunately, my gynecologist did a biopsy which showed carcinoma in situ, and it was treated before I might have become a cervical cancer statistic.

We therefore urge the Committee, if review of the materials warrant, to make available to clinicians and women more options for increased accuracy in Pap smear testing, in particular, to review whether automated techniques may be useful for primary screening, as well as for quality

control.

Thank you.

DR. FRANCIS: Thank you, Ms. Anderson. I would like now to ask Diana Silman, Silver Spring, Maryland to address the Panel. Please remember to state whether you have any financial involvement with the sponsor or their competitors.

MS. SILMAN: Good morning. I do have no financial interest in NeoPath or any other company. I am not being compensated by anyone. None of my expenses are being paid for my participating in this hearing. I am not a doctor, I am not a nurse, I am not a director of a health center. I do not know the first thing about the medical field, except that I, too, am a I survivor of cervical cancer, and needless to say, very nervous, and I feel very strongly about this issue.

I was diagnosed in March of 1990, but only after six months of confusing and inconclusive tests, including a number of false negatives. I was lucky, only because my mother, a cancer survivor herself, was overly cautious.

She knew that I was high risk, and she knew that the whole process involving Pap smears and screening for cervical cancer left a lot to be desired. And she was right.

One month after a false negative, my Pap test came back Class III, positive for cervical cancer. I was under the knife within a week. Although the doctors recommended a hysterectomy, I opted for laser cone surgery; instead of removing my entire reproductive system at the age of 27, they removed a cone-shaped chunk of my cervix; this way, I figured I could still keep hope alive of having a family one day, but unfortunately, I have been unable to carry a pregnancy to term. And I feel that all of this could have been prevented.

There needs to be a better screening process for this disease, period. Approximately 5,000 women, as we know, lose their lives every year to cervical cancer, and that is 5,000 too many. This is a curable disease. When diagnosed early and accurately, it is curable. And like I say, I am not a doctor, but I know this. It is curable. I am one of them. I was lucky enough. I am living proof. And I am here to help put a face to some of the statistics that we hear.

I am a cancer survivor, but my relationship with this disease will go on for a lifetime. Every three months, I have a Pap smear, and every three months I wait for those results on pins and needles. And when it comes back negative, I have to wonder; is that accurate?

If there is any technology out there that can help decrease the number of false negatives, decrease the incidence of human error, and increase the number of tests that one lab can do, then in my book, this is a no-brainer.

In my book, it will save lives. Put a face to those statistics, not just mine; not just Carol Ann Armenti's, but think of yourselves, think of your own families; think of your mothers, think of your sisters, and think of your daughters.

Thank you.

DR. FRANCIS: Thank you, Ms. Silman. Is Dr. Stanley in the audience, yet? In that case, I would like to call upon Dr. Stanley, Chief of Pathology at Hennipin County Medical Center to address the Panel.

Dr. Stanley, if you would proceed your comments by stating whether or not you have any financial involvement with the manufacturers of any products being discussed, or their competitors, please?

DR. STANLEY: I do not. Mr. Chairman and members of the Panel, my name is Michael Stanley, MD. I am Chief of Pathology at the Hennipin County Medical Center in Minneapolis, Minnesota, and I am Professor in the Department of Pathology at the University of Minnesota.

I am here today representing the American Society

of Clinical Pathologists. I serve on its Continuing Education Council for Cytopathology. The ASCP is a nonprofit medical specialty society, organized for educational and scientific purposes. Its 75,000 members include board-certified pathologists, other physicians, clinical scientists, and a variety of certified laboratory technologists and technicians.

Automation in cytopathology is not new, however, its use in the clinical set is a relatively recent development. It is for this reason that the FDA -- I am sorry, the ASCP -- urges the FDA to carefully consider standards for primary screening instruments for gynecologic cytology, already developed and published by the Intersociety Working Group for Cytology Technologies.

Let me briefly address some of the function of these instruments, initially. Since primary screeners are devices that are intended to triage gynecologic cytology slides for identification of malignant and premalignant lesions, and atypical squamous cells of various types, it is essential that even low level abnormalities, including those currently designated as being of uncertain diagnostic significance, should not be excluded from review.

Furthermore, devices should have the capability to comment electronically on things such as specimen adequacy,

the presence of an endocervical component -- the importance of which I will be pleased to discuss further at your pleasure -- or, the presence or absence of infectious agents, which are very important in certain patient populations.

If the device lacks these capabilities, then each slide should be reviewed manually for these components. However, it should be noted, and it has been published, that the type of cursory review involved in this consideration is not tantamount to the sort of screening that traditionally goes on in the search for malignant or premalignant conditions.

There are two functional categories for primary screening devices. A device may be interactive, and in that case, a computer identifies and then presents to a human potentially abnormal cells, which a cytotechnologist reviews, and then decides whether or not a slide should go on for further human review.

Alternatively, a device may be independent; in that case, a computerized slide is examined by the machine; a score is assigned to that slide, based on whatever degree of abnormality is noted; and if its score falls below a certain preset threshold, that slide is then excluded from further human evaluation.

Primary screeners that function interactively require both high quality image presentation and appropriate training for cytotechnologists and pathologists to evaluate the images in all applicable practice settings.

Likewise, primary screeners that function independently should pass no slides that later show -- or turn out to be -- representative of, high grade lesions or invasive cancers, because we cannot afford to miss these cancers.

A few comments about sensitivity, if I may. When reviewing these primary screening devices, the sensitivity of the instrument must be considered carefully. The sensitivity of the device, plus the expertise of the cytotechnologist, should equal or exceed or the sensitivity of primary manual screening as it currently exists and is practiced.

Performance specifications including sensitivity and positive predictive value of the instrument alone, and in combination with manual screening, should be determined in a prospective, blinded, adjudicated trial, as noted more fully in a moment.

It is critical that performance specifications be noted separately for cells currently designated as being of low level abnormality, including those of uncertain

diagnostic significance, and then separately for those that are of low grade intra-epithelial lesion, or of high grade intra-epithelial lesions plus cancer.

This is especially so, since it is difficult to achieve expert consensus on what constitutes ASCUS -- that is, uncertain significant cells -- and what the significance of those abnormalities may be to many of the women in whom they are identified. These specifications should be made available to potential users, and the results must be reproducible.

Device algorithms may need improvement, and study design may need reevaluation, if the primary increases in sensitivity turn out to be largely confined to identification of more cases with cells that are abnormal, but at the level of uncertain diagnostic significance, rather than detecting more cells that have significant diagnostic abnormalities.

Furthermore, the device should identify for human review the full range of glandular cell abnormalities in the group now collectively known as glandular cells of uncertain diagnostic significance, a task which I would like to comment, is currently deemed very difficult for even the most experienced human observers by many experts.

Device manufacturers should make a number of

disclosures to those who might seek to use these technologies. These should include the effects of stain variability on the results, not only from lab to lab, but within a lab, a given lab, from day to day or week to week.

Whether or not the machines can identify an endocervical, endosquamous component, important determinants of the adequacy of the slide in the first place. The identification of unsatisfactory slides is also very important. Those patients need other evaluation.

The identification of endometrial cells, although not the primary thing we look for on cervical smears, and although shed from a different part of the uterus, these cells are of critical diagnostic significance in certain patient populations, particularly postmenopausal women. Their identification is important; limitations in their identification should be disclosed, if they exist.

Other disclosures include the ability to identify infectious organisms, the acceptance of slides produced by a range of preparatory methods and devices, and also, the slide rejection rate, in terms of the number and types of cases not suitable for machine evaluation.

Regarding advertising of such devices, the Intersociety standards state specifically that manufacturers should refrain from claiming new standards of practice in

their product advertising. Professional standards are set by professionals over time, not by advertisers and manufacturers.

Let me briefly run through some study design requirements that we feel are needed to generate data adequate for device approval.

To determine performance acceptability, a study of primary screening devices must be compared to contemporary primary manual screening for the same slide population. Contrary study results from the two study arms should then be adjudicated by an independent Panel of experienced cytology professionals. Consensus should be achieved; however, it is important to point out that consensus, in study arm differences, may be very difficult to achieve, and must meet some definition that is approved by the profession.

When evaluating machine performance, adjudicated cytology should be used as the gold standard. This is to account for other problems. For example, a positive biopsy and a negative cytology could indicate cytology sampling or preparatory errors.

A negative contemporary biopsy with positive cytology may indicate sampling error at the time of biopsy. Both of these problems are real; both of them occur on

virtual daily basis in the practice of gynecologic cytology and pathology.

Finally, we should note that a positive biopsy taken more than six months after a negative cytology may represent the development of a new lesion that has nothing to do with the initial cytology.

The study design should also consider positive predictive values. Tissue biopsy of a statistically significant subset of patients with positive cytologic diagnoses should be used as the gold standard for this determination.

A comparison of the receiver operator curve characteristics for the two study arms, whether cytologic or histologic, should be considered.

It is also important to eliminate bias resulting from differences in cytotechnologists' experience and abilities. Cytotechnologists examining device performance must be similar in training and experience to those in the human primary screening arm of device evaluation studies, rather than being specially trained experts. In short, the study should replicate intended use conditions as closely as possible.

The Intersociety Working Document and other recent publications address a great number of other issues which we

do not have time to discuss this morning, including the potential increased cost for computerized testing with subsequent effects on the public health, medical legal implications, and even a possible increase in the number of slides missed by implementation of devices that function within certain ranges of performance parameters. Thus, we urge careful consideration by the FDA of the entire Intersociety Working Group document.

In concluding, I would like to say that adhering to the standards produced by the Intersociety Working Group for Cytology Technologies, of which the ASC is a participant and signatory is critical to assuring that new primary screening devices meet or exceed existing standards for primary manual screening.

My final statement will be the most appropriate of all, I think. For the benefit of the public health, the ASCP wishes to reiterate that women should continue to obtain Pap smears at regular intervals as established by their personal physicians, no matter what the method of screening in the laboratory.

Thank you.

DR. FRANCIS: Thank you, Dr. Stanley. I would now like to read into the record three statements submitted to the Panel.

The first was submitted on behalf of the American Social Health Association, ASHA, as testimony for consideration by this Panel.

The American Social Health Association, or ASHA, is a non-profit organization established in 1914 to eliminate sexually transmitted diseases and their harmful impacts on individuals, families, and communities. Because of the link between certain types of the human papillomavirus and cervical cancer, ASHA has a significant interest in Pap smear screening.

ASHA sponsors a number of consumer programs dedicated to the prevention, early detection, and control of cervical cancer. For example, ASHA operates the CDC National STD hotline, and provides information directly to consumers about cervical cancer and Pap smear screening.

ASHA also sponsors support groups for women with HPV, focusing in part on the importance of regular Pap smears. Thus, ASHA represents hundreds of thousands of women who are concerned about Pap smear screening and cervical cancer.

Despite the success of the Pap smear, American women still die from cervical cancer. The American Cancer Society estimates that 5,000 women will die in 1998 from cervical cancer. Yet, all of these deaths can and should be

prevented.

Multiple factors contribute to this problem. These problems include barriers to screening, and false negative Pap smears. ASHA encourages the development of new technologies to improve the quality of Pap smear screening.

ASHA also hopes that these new technologies will help ensure that all women have access to cervical cancer screening. As organizations, such as ASHA, encourage more women to participate in Pap smear screening, our health care system must be able to handle the increased workload, without sacrificing quality.

If we cannot achieve this goal, we have failed. ASHA hopes through the combined efforts of health care professionals, industry and FDA, the quality and availability of Pap smear screening will continue to improve, and that all women will have the opportunity to benefit.

The next statement was submitted by the Trylon Corporation. Dear Ms. Calvin. I am taking this opportunity to write to you regarding the upcoming Hematology and Pathology Devices Panel Meeting on January 28, 1998.

This letter is being submitted to be officially read into the record during this Panel meeting. I am requesting that the Panel also consider the fact that the

terminology used to describe the Neopath AutoPap System be revised to remove the terminology primary screening.

The system is clearly a Pap smear reading device, designed to scan Pap smear slides and use a computer algorithm to determine which slides contain abnormal cells and which slides do not.

The new data submitted may show that when this AutoPap System is used to evaluate slides, it was superior to the Current Practice arm for identifying the 25 percent of slides that contained no abnormal cells. This may well represent an advantage worthy of raised claims, but these data are a far cry from establishing the AutoPap Device as a primary screening device.

We have made this same argument in the past regarding the claims that have already been approved for in-vitro devices governed by the Hematology and Pathology Devices Panel. This terminology request surfaced in July, 1997 when the Obstetrics and Gynecology Devices Panel correctly decided that cervical pathology, as determined by biopsy, was the appropriate gold standard to use when measuring the efficacy of proposed screening adjuncts.

Our FDA applications for the Pap Plus Speculoscopy system have had to meet this metric, and we have recently been granted claims from FDA that Pap Plus Speculoscopy is a

more sensitive primary screening test for cervical pathology of LSIL+ than the Pap smear alone.

If biopsy is the gold standard that a screening test for cervical disease is to be measured against for the Obstetrics and Gynecology Devices side of the FDA, then inconsistencies in the Hematology and Pathology branch must be corrected.

On the one hand, the public and providers are being told that FDA recognizes the importance of establishing a tissue diagnosis against which to measure a screening test, and on the other hand FDA is allowing the terminology screening to be used for a cytology evaluation device that uses other cytological results as the gold standard. The Pap Plus speculoscopy claim clearly shows that the Pap smear alone, when measured against tissue biopsy, has a sensitivity of less than 50 percent!

Although this terminology issue may not seem significant, owing to the fact that for many years we've referred to the Pap smear as a screening test, the data are overwhelming that cytological testing does not have a high sensitivity when biopsy is the gold standard.

FDA has data in hand to demonstrate this, and with new in-vivo tests being developed, will have much more data along the same lines in the future. It is important that

FDA not mis-inform providers and patients regarding the tests they clear for use, and it is our opinion that at every opportunity, such as this opportunity for Neopath's AutoPap System, FDA take the time to correct the language of the claims it approves.

The AutoPap System is an evaluation tool for reading Pap smear slides. It may be viewed as a primary reading tool if FDA agrees with this claim for Neopath, but it is not a primary screening device. Even the term primary Pap smear screening device is confusing, since the word screening evokes the idea that a screening test has been conducted, when in reality when a Pap test is performed only an evaluation of exfoliated cervical and vaginal cells has taken place.

In years gone by, we may have believed that this was screening, and it may well have been that an evaluation of these exfoliated cells was the best screening we could perform. However, at this stage, FDA is well aware of the data that show that cytology alone may be so insensitive to cervical pathology, that the term screening test is inappropriate.

We urge the Panel to consider the terminology it uses. There is enough mis-information in the public sector already, and numerous editorials have been written about the

unrealistic expectations of a public who thinks no case of cervical cancer should be missed because the Pap smear is so highly touted as a screening test.

The litigious nature of the missed cancer diagnosis arena is well known to even those with just a passing interest in this field. The expectations of the public from claims made regarding the Pap smear as a screening test, are clearly out of proportion to the test's capabilities.

FDA has an opportunity, and we would suggest an obligation, to do everything it can to clear up these muddied waters.

The letter is signed, Stewart A. Lonky, MD, FACP, Medical Director, the Trylon Corporation.

The last statement I will read is being submitted by the Society for the Advancement of Women's Health Research.

The Society for the Advancement of Women's Health Research (SAWHR) is a non-profit organization established in 1990 to improve the health of women through research.

SAWHR's mission is to increase the funding of research dedicated to the prevention and treatment of diseases, and conditions prevalent among, and disproportionately affecting, women.

SAWHR is particularly interested in cervical cancer screening, and how new technologies may improve the quality of screening. According to the American Cancer Society, cervical cancer rates have dropped by 70 percent since the introduction of the Pap smear. Despite this success, it is estimated that 5,000 women will die of cervical cancer in 1998.

This figure does not include the tens of thousands of women who will undergo traumatic therapies to treat cervical cancer and its precursors. The conventional Pap smear has limitations which may lead to false negatives. The false negative rate may be as high as 25% in some laboratories. Sampling errors, preparation of the smear, and screening errors may all result in false negative readings.

New technologies are needed to minimize these errors, and to improve the overall quality of cervical cancer screening. These advancements should be made available to women and their health care providers as soon as their benefits have been demonstrated in appropriate clinical trials.

SAWHR encourages this Panel and FDA to fairly evaluate the benefits and risks of new technologies, such as the application being considered today. Companies, such as

NeoPath Inc., must be encouraged to continue their research and development of new technologies. It is only through the collective efforts of industry, health care professionals and FDA that the overall quality of Pap smear screening will improve, and more women will benefit from successful screening.

That concludes the written statements. We are a little over time, so I will curtail any discussion on the presentations, and move right ahead to presentations by the sponsors, NeoPath, Inc.

The first presentation will be by Mary K. Norton, Director, Regulatory and clinical affairs.

Agenda Item: Sponsor Presentation -- NeoPath, Inc.

MS. NORTON: Well, good morning. My name is Mary K. Norton. I am the Director of Regulatory and Clinical Affairs at NeoPath, Incorporated. I will be making NeoPath's introductory remarks, and I will see to it that your questions are answered to the best of our abilities.

On behalf of NeoPath, I want to thank each of you for your contributions to the review of this application. During the next few minutes, I will provide background information about the healthcare problem we are addressing and remind you of the history of our device. Also, I will

describe the reasons why the AutoPap System should be approved.

Since its inception in 1989, NeoPath has been committed to decreasing the impact of cervical cancer by improving the methods of early detection. About 500,000 new cases of cervical cancer and its precursor conditions are reported world-wide each year. In the United States, approximately 15,000 new cases will be diagnosed in 1998, and approximately 5,000 women will die of the disease.

Early detection of precursor lesions can reduce, for hundreds of thousands of women, the trauma associated with the corrective therapies and the risk of cervical cancer. The screening of Pap smears is the standard procedure to detect pre-cancerous and cancerous conditions of the cervix. Despite the acknowledged diagnostic value of the Pap smear, there is a high false negative rate. In some laboratories, this rate may exceed 25 percent.

To lessen this problem, the Clinical Laboratory Improvement Amendments of 1988 directed laboratories to conduct a number of quality control procedures, including rescreening a minimum of 10 percent of all Pap smears initially classified as normal.

This 10 percent quality control rescreening requirement helped monitor laboratory proficiency and did

somewhat reduce the false negative rates. The improvement in false negative rates, however, was expectedly small. Therefore, CLIA '88 mandated quality control rescreening is not the complete answer to the problem.

To address the problem of false negatives, NeoPath began exploring technological solutions and developed an automated slide analysis system known as the AutoPap System. Our early studies compared the AutoPap's performance as a quality control rescreening device to the existing 10 percent random selection process.

These early studies, which included a large number of high grade lesions and cancer slides, demonstrated the AutoPap's high sensitivity to cancerous and pre-cancerous conditions. These studies also showed that the AutoPap provided a five- to eight-fold improvement in the recovery of false negative slides. As a result, this Advisory Panel recommended approval, and the FDA approved the AutoPap 300 QC System in 1995.

Despite this advancement, we recognized that improving quality control alone was not enough. Thus, NeoPath began to evaluate the AutoPap's use as a combined initial screening and quality control rescreening device.

In 1996, we submitted a supplemental application to use the AutoPap System in the initial screening process.

The submission was based upon data obtained from an expanded analysis of the earlier studies. At that time, this Panel recommended that we obtain additional clinical information to show that the AutoPap was as good as, or better than, current practice in an actual laboratory environment, and the FDA agreed.

The FDA and the Panel requested NeoPath to collect this information in a new, intended use prospective study, which was to include a sufficient number of slides, particularly for LSIL, and an appropriate truth determination plan.

We developed a new study protocol, conducted the study requested by the FDA and the Panel, and submitted the supplemental application you are considering today. We also took this opportunity to upgrade the AutoPap with an additional algorithm to improve the detection of glandular abnormalities and unsatisfactory samples.

The following are the reasons why the AutoPap System should be approved for commercial distribution without delay.

First. NeoPath has conducted a prospective, masked, matched controlled study entirely consistent with prior FDA and Panel advice, and with the recommendations of the Intersociety Working Group.

Second. The NeoPath study provides clinically and statistically significant and unequivocal results.

Third. The NeoPath study was conducted in laboratories that operate under current laboratory practice standards. For this reason, other laboratories will experience the same enhanced performance with the AutoPap System.

Fourth. The AutoPap is designed for use with conventional slide preparation techniques used by the vast majority of laboratories in the United States. Thus, it is a technological enhancement that will have practical application throughout this country.

Fifth, and most importantly, the AutoPap will assist laboratories in reducing false negatives. It will help save lives. It will significantly advance health care for women.

The new AutoPap improves laboratory performance in three ways. First, the AutoPap helps cytotechnologists identify more false negatives during quality control rescreening. However, we know that improving quality control rescreening alone is not enough -- in part, because the purchase of the quality control rescreening device adds to overall laboratory costs. Thus, secondly, the initial screening of Pap smears by the AutoPap allows laboratories

to process more slides, which compensates for the cost associated with the purchase of the device, and to detect even more abnormalities.

Finally, the new AutoPap reduces the number of false positive slides. The AutoPap screens and identifies a set of slides most likely to contain abnormal cells, and ranks these in order of probable abnormality. These slides are screened by cytotechnologists using the ranked review report.

The AutoPap also identifies a set of slides for which the probability of abnormality is very low. These slides are screened by only the AutoPap. When the AutoPap is used in today's laboratory, the cytotechnologist is expected to screen as many slides as in the past. Thus, by eliminating certain slides from the cytotechnologists' workload, AutoPap initial screening enables the laboratory to process more slides.

We would like to state that NeoPath fully understands the role of the Panel, and of the FDA. We know you are here to assess the safety and effectiveness of our device, not its cost-effectiveness. Indeed, our presentation will emphasize that the AutoPap is safe and effective, only because our study has conclusively shown that it helps cytotechnologists identify abnormal slides.

Nevertheless, without designing a device that can be purchased and used within the financial constraints of our health care system, we cannot deliver the benefit of increased performance. For this device to fit within our health care system, it must assist the laboratory in both the initial screening and quality control rescreening of Pap smears.

Among future plans is a further enhanced AutoPap System. For example, we will be studying the effect of providing more information to the cytotechnologist, such as the locations of cells of interest and information about why the AutoPap has assigned a particular ranking to a slide. We believe this will further improve sensitivity.

I mention this to put aside questions about whether future developments are important in today's deliberations. The AutoPap you are considering today is superior to current practice. Unlike other options, the AutoPap is usable by the vast majority of laboratories in the United States and it is available for immediate distribution.

For these reasons, the AutoPap should be approved without delay. In our presentation, Dr. Alan Nelson, our President and Chief Executive Officer, will describe the AutoPap's theory of operation and quality assurance.

Dr. David Wilbur, Medical Director for NeoPath, Associate Professor of Pathology and Laboratory Medicine, and Co-Director of the Cytopathology Laboratory at the University of Rochester Medical Center will describe the design of our study and the study results.

Finally, Dr. Thomas Bonfiglio, Senior Attending Pathologist and Director of Cytology at The Genesee Hospital, and Clinical Professor of Pathology and Laboratory Medicine at the University of Rochester, will summarize the reasons for the approvability of the AutoPap.

All speakers and Dr. Richard Chiacchierini, our statistical consultant, will be available to answer questions during and after our presentation. In addition, several of our clinical investigators, including Dr. William Tench, Palomar Medical Center, and Dr. Marianne Prey, Smith Kline Beecham Clinical Laboratory, St. Louis, are in attendance and will be available to answer questions.

Thank you very much for your attention. I would now like to introduce Dr. Alan Nelson.

DR. FRANCIS: Just for the members of the Panel, we shall be taking questions and comments at this time at the end of these presentations.

DR. NELSON: Good morning distinguished Chairman and members of the Panel. I am Alan Nelson, NeoPath's

founder, and its President and CEO. I am a physicist by training and formerly was a professor of Nuclear Engineering at MIT, where I directed the joint Radiological Sciences Program between MIT and Harvard.

Later, as a professor of Bioengineering, I directed the Center for Imaging Systems Optimization and the Medical Imaging Graduate Program at the University of Washington.

I founded NeoPath nearly ten years ago with one purpose: to improve the accuracy and availability of Pap smear screening through the use of automation to help eradicate cervical cancer.

I will describe our product, the AutoPap System, by discussing its theory of operation and its quality assurance process. The AutoPap System consists of an optical-mechanical unit that positions Pap smear slides and scans them, using two magnifications, to gather cell image data.

The cell image data are fed to a computer, which measures morphology and calculates various cell parameters. The computer runs three image interpretation algorithms that analyze every cell to determine normality and abnormality. These algorithms are the core technology of the AutoPap System, so I want to explain how they work.

The new AutoPap, which you are considering today, contains two fundamental algorithms, which were independently trained and developed. One designed to optimally search for squamous cell abnormality, and the other for glandular cell abnormality.

These two algorithms work in parallel, analyzing hundreds of critical features of cells, and groups of cells, to ultimately assign a score to each slide. This score is a measure of likelihood that the slide is normal or abnormal.

Each algorithm independently assigns a score, but the final score assigned to the slide is the most conservative of the two scores. The AutoPap uses the final score to rank order slides according to likelihood of abnormality.

The slides that receive the highest ranks, those most likely to be normal, are reviewed by the AutoPap only and are classified within normal limits.

Slides with the lowest ranks, those more likely to be abnormal, will certainly receive another review by a cytotechnologist, and if presumed to be normal, will be selected again for quality control rescreening.

Here is a curve showing how the AutoPap ranks a population of slides. As I mentioned, the AutoPap scores are converted into a rank, where high ranks represent the

high scores. A distribution of slides -- approximately 95 percent of a population -- would be so ranked, and the smaller portion -- approximately 5 percent of slides -- would receive the lower scores in the AutoPap System.

NeoPath's original primary screener submission, which this Panel reviewed approximately 15 months ago, was based on an older AutoPap with only one algorithm that, in a typical population of slides, would have selected no more than 30 percent of slides as AutoPap Review Only, and 10 percent of slides as probably abnormal, to receive quality control rescreening.

The submission before you today is based on the latest generation AutoPap with two algorithms. The second algorithm adds conservatism to the system by removing more slides from the AutoPap Review Only population, and adding more slides to the quality control rescreening population, as indicated by this arrow, and this arrow. This arrow, by the way, is the direction of conservatism in reading slides.

This results in no more than 25 percent of slides selected for AutoPap Review Only and at least 15 percent selected for quality control rescreening. Because every slide is scored independently, when there is a high prevalence of disease in a group of slides, it is entirely

possible that no slides will be designated as AutoPap Review Only, and more slides may be designated for quality control rescreening.

In this submission, we include previous data from studies with retrospective and prospective slides to demonstrate AutoPap's exceedingly high sensitivity to abnormal slides, particularly high grade lesions and cancer, which could not possibly be obtained in any realistic prospective study.

These studies are referred to as the Historical Sensitivity Study and the Current Archive Study. These sensitivity data, combined with other clinical studies, resulted in our first FDA approval to use the AutoPap System as a quality control rescreener.

This approval was based on the older AutoPap with only one algorithm for scoring slides and establishes the device's minimal performance. The data from the prospective, intended use study before you today demonstrate that the AutoPap with two algorithms has achieved an accuracy that now warrants approval as an effective primary screener of Pap smears.

To ensure the AutoPap's consistent performance and compatibility with different specimen preparation among laboratories, NeoPath has implemented a three-fold quality

assurance program, comprised of laboratory process compatibility -- what we call LPCA -- that keeps continuous track of differences in laboratories as we install systems.

Two, laboratory process monitoring, which looks at any potential changes in the laboratory process, particularly slide preparation and staining.

Finally, system integrity, which keeps track of the AutoPap's internal parameters to ensure the AutoPap is always working correctly.

Once the adjustments are set by NeoPath personnel for LPCA -- the overall performance -- the parameters remain constant and cannot be changed by the laboratory; however, to detect changes that may occur over time in the laboratory slide preparation process, the AutoPap continuously monitors this, using the Laboratory Process Monitoring Procedure.

If the laboratory slide preparation does change, such that the AutoPap approaches its tolerance limits, the AutoPap will simply stop processing slides, until NeoPath service personnel can rerun LPCA to accommodate the new slide preparation process. The AutoPap's tolerance limits are set to ensure accuracy.

Finally, the AutoPap has an internal automated quality assurance program which checks and confirms that each subsystem is operating within specifications. We call

this program System Integrity. Before each slide tray is processed, the System Integrity verifies that the data collection and image analysis algorithms are operating within specification. It confirms that the video microscope is in proper alignment. It calibrates both the low and high power magnification.

If a problem arises that might compromise the system's accuracy, the AutoPap alerts the operator, and rejects the tray. NeoPath's customer service is called to resolve the issue.

Taken together, Laboratory Process Compatibility Assessment, Laboratory Process Monitoring, and System Integrity provide an overall quality assurance and confidence that the AutoPap reports accurate and reliable slide information.

We are proud of the AutoPap System's accuracy and reliability under continuous operation. The AutoPap quality control product has been used by U.S. laboratories for nearly two years, and its technology is proven.

I would now like to introduce Dr. David Wilbur, who will review the clinical data to support the specific claims of using the AutoPap as a primary screener.

Thank you.

DR. WILBUR: Thank you, Dr. Nelson. Good morning

distinguished Chairman and Panel members. I am Dr. Dave Wilbur, the Medical Director for NeoPath. I am also a board-certified Cytopathologist and Laboratory Director with 14 years experience in clinical cytopathology.

It is my pleasure today to present you with the data from the new clinical trial of the AutoPap System. The clinical trial includes over 25,100 Pap smear slides. This study adds to the performance results of previous clinical studies, which demonstrated the AutoPap's sensitivity to precancerous and cancerous slides and its benefits in quality control rescreening.

I will briefly discuss the study design, how the AutoPap works in the laboratory, and the clinical results, including the sensitivity and specificity improvements. In every way, the study conforms to the requirements that were defined by the FDA, members of this Advisory Panel, and with the recently published guidelines for the Intersociety Working Group.

It is a prospective study with appropriate statistical design, which evaluates the slides in accordance with the diagnostic and adequacy categories of the Bethesda System. It directly compares the laboratories with and without AutoPap in actual use conditions with comprehensive truth adjudication of discordant cases by an independent

Panel of cytopathologists.

The objective of this study was to test whether the AutoPap-assisted practice -- which I will describe in a few minutes -- was as good as, or better than clinical laboratory current practice. The study was prospective and designed to emulate intended use conditions.

Five cytology laboratories, covering a broad demographic range of women, participated in the study. Over 25,100 slides were fully analyzed and subjected to comprehensive truth adjudication.

Each cytotechnologist screened slides on both study arms, and all were masked to any prior results. However, at no time did the same cytotechnologist review the same slide between the study arms. Nearly all cytotechnologists from each laboratory participated in the study and were required to review slides while meeting their normal workload requirements.

This study compared AutoPap-assisted practice with current practice. Current practice consists of the 1988 CLIA-mandated 100 percent manual, initial screening and random 10 percent manual quality control rescreening.

The AutoPap-assisted practice consists of 100 percent AutoPap initial screening of Pap smears, followed by AutoPap-assisted manual screening of approximately 75

percent of the slides; and finally, AutoPap-assisted manual quality control rescreening of at least 15 percent of the slides originally classified as within normal limits.

Collectively, we refer to these three processes as the AutoPap-assisted practice. I would now like to provide an overview of the AutoPap-assisted work flow, or how the AutoPap fits into the laboratory.

All slides are entered into screening, with the exclusion of high risk slides; 100 percent of slides therefore after these exclusions are processed on the AutoPap System.

Following this initial 100 percent AutoPap screening, the AutoPap identifies a subset of slides, at least 75 percent -- and I will describe what happens to the other slides in just a moment -- these 75 percent of slides that are most likely to contain the abnormal cases.

Each slide in this subset is ranked according to the likelihood of abnormality. The AutoPap's ranking of each slide is provided in a report to the cytotechnologist, who then reviews the slides using the ranked review report. The ranked review report provides three important pieces of information.

The first is the AutoPap's ranking of each slide for probable abnormalities. Each slide is individually

ranked from 1 to 100, where a rank of 1 indicates slides with the highest probability, or highest likelihood, of abnormality. In addition, the slide is assigned a slide group ranking, ranging from 1 to 5, where a rank of 1 indicates the group with the highest likelihood of abnormality.

The second piece of information on the report is the AutoPap's evaluation of slide adequacy according to the Bethesda System. Each slide is classified as satisfactory, satisfactory but limited by, or unsatisfactory. The presence of squamous component and endocervical component is reported as detected or not detected, and inflammation and/or obscuration is reported as a percentage of the slide coverslip area.

Finally, the report confirms that the slide was completely and successfully processed by the AutoPap. The cytotechnologist will determine whether the slide is within normal limits, unsatisfactory, or potentially abnormal.

As in current practice, potentially abnormal slides are forwarded to a pathologist and manually screened. Within normal limits slides receive no further evaluation, except that they are eligible for quality control rescreening.

All slides that are classified as within normal

limit by the cytotechnologist following manual review -- this arm, as illustrated here -- are eligible for quality control rescreening to identify manual screening false negatives. This is illustrated in the box in the lower portion of the field.

The laboratory again uses the AutoPap's ranking information as a guide to select approximately 15 percent of these presumed normal slides, for manual quality control rescreening by a cytotechnologist.

These are the slides most likely to include the false negatives. Following this review, slides are signed out as within normal limits, as unsatisfactory, or are sent on to the pathologist as potentially abnormal.

As I mentioned previously, and as shown here on the left side of the illustration, the AutoPap identifies a subset of slides, up to 25 percent of the total, that have the highest probability of being normal.

These slides are screened by only the AutoPap System and are classified as within normal limits. A report which provides adequacy information is also generated for these slides.

Now, I would like to move on to discuss the results of the clinical study. As shown here, the study evaluated the performance of both practices on three

principle disease categories.

The first is ASCUS+, which includes all abnormal cases; the category of LSIL considered alone; and the category of LSIL and greater, which includes low grade squamous intra-epithelial lesion, high grade SIL, adenocarcinoma in situ, and all cases of cancer. And in addition, two adequacy categories were considered: Satisfactory but limited by, or SBLB, and Unsatisfactory.

NeoPath chose these categories, with the agreement of the FDA and this Panel, to achieve a balance between clinical significance and the practical realities of conducting a prospective study.

We chose ASCUS+ to demonstrate the performance of the AutoPap on all abnormal. And since we know that the clinical significance of ASCUS and AGUS may not be clear for all women, we also looked at the device's performance on LSIL alone, and on LSIL+, since we know absolutely that these categories are clinically important.

The study also evaluated the performance of both practices on specimen adequacy, in accordance with The Bethesda System categories of SBLB and Unsatisfactory. Any initial screening system must identify slides that are unsatisfactory for interpretation.

Women with unsatisfactory Pap smears may be at

higher risk for having an undetected abnormality and need to be identified so that they can be appropriately followed.

To compare the two study arms, we also must know slide truth. Our truth adjudication process covered both slide diagnoses and adequacy determinations, and followed the recommendations of both the FDA and members of this Advisory Panel.

It also complied with the protocol described in the Intersociety Working Group guidelines. Remember that, when screening diagnoses were concordant -- that is, agreement between the study arms -- this was considered to be truth. When screening diagnoses were discordant, an external discrepancy Panel was convened to determine cytologic truth.

The external discrepancy Panel consisted of groups of three cytopathologists who independently diagnosed a slide. A truth diagnosis was determined if two out of three agreed; otherwise, the slide was reviewed at a multi-head microscope until a consensus diagnosis was achieved.

When adequacy determinations were concordant -- or when the two arms agreed -- this was also considered to be truth. When the adequacy determinations were discordant, a single, independent senior cytotechnologist reviewed the slide to determine truth. In both discrepancy

resolution processes, all observers were masked to all the prior diagnoses.

To meet our objectives, we determined that the study required a very large number of slides. A total of 31,507 slides entered the study across all five laboratories; 5,336 were ineligible according to the inclusion and exclusion criteria of the study. Of these, 3,200 were high risk slides.

Each laboratory used its current definition of a high risk slide. The other slides were excluded for suitability reasons that included damaged slides, slides with excessive air bubbles, or slides that were part of a multi-slide case. Of the remaining 26,171 slides submitted to the study, only 3.68 percent -- or 963 slides -- failed to process on the AutoPap, generally due to problems with the slide, the coverslip, or the preparation of the specimen.

This left 25,208 slides of which 25,124 -- the number at the bottom right -- had truth and were available for final analysis. This slide shows, by the Bethesda System category, the total number of slides analyzed.

There were 171 Unsatisfactory slides. The within normal total included 5,873 Satisfactory But Limited By slides. Of the 1,397 abnormal slides, there were 998 cases

of ASCUS, 51 cases of AGUS, 278 LSILs, 67 HSILs, 1 AIS, and 2 cancer slides.

This slide shows the results for all abnormal slides in the study, which includes ASCUS, AGUS, LSIL, HSIL, AIS, and cancer, a total of 1,397 slides, as you can see here, illustrated in the lower right-hand portion of this 2 x 2 matrix.

In this table, the rows across represent the results from the AutoPap-assisted practice, and the columns up and down represent the results from current practice. You can see that 908 abnormalities -- illustrated in the upper left-hand box -- were detected in both study arms.

The AutoPap-assisted practice, however, detected another 291 cases -- in the upper right box. These were abnormal slides that current practice missed, so, the AutoPap-assisted practice detected a total of 1,199 slides, or abnormal slides. In contrast, current practice detected, in addition to these 908 slides, 198 abnormal slides that the AutoPap-assisted practice missed, for a total of 1,106 abnormal slides.

The box on the lower right -- illustrated with the dash -- represents the slides that were classified as within normal limits by both study arms, and is left blank because normal slides are not considered for the analysis of

performance on abnormal slides.

To determine whether these numbers -- 291 and 198 -- were statistically the same or different, we used a conditional binomial test, or equivalence test. This is the appropriate statistical test for this study design.

In this analysis, we set the p-value at the standard level of 0.05, to determine whether the AutoPap-assisted practice is at least as good as current practice.

Since our p-value as illustrated here -- .00 -- is much lower than 0.05, there is no question that we are least as good as current practice.

We then used a different binomial test, or superiority test, to test for statistical superiority. Our p-value -- again, illustrated on the bottom of the slide at 0.0 -- for statistical superiority, demonstrates conclusively that the AutoPap-assisted practice is clinically and statistically superior to the current practice in assessing ASCUS+, or all abnormal cases.

This slide shows the results for truth adjudicated LSIL slides only. When comparing the two study arms, you see that the AutoPap-assisted practice correctly detected 253 slides, which is greater than the 233 slides that were detected by current practice.

Applying the equivalence and superiority tests,

the p-values -- illustrated at the bottom -- show that the AutoPap-assisted practice is at least equivalent -- by the top p-value -- and it is in fact, superior to current practice -- by the bottom p-value -- for LSIL considered alone.

This slide shows the results for truth adjudicated LSIL+ cases, which again includes LSIL, HSIL, AIS, and cancer. The AutoPap-assisted practice correctly detected 321 slides, compared to 298 slides detected by current practice.

The 321 slides correctly called abnormal by the AutoPap included two cancer slides, both of which current practice missed. Again, applying the equivalence and superiority tests, the p-values show -- as illustrated here -- that the AutoPap-assisted practice is at least equivalent with this p-value, and is in fact, statistically superior to current practice -- as illustrated by the bottom p-value -- for categories of LSIL and above.

In summary, the AutoPap-assisted practice detected more abnormal slides in the LSIL and HSIL categories. The false negatives were reduced in virtually all diagnostic categories from ASCUS to cancer, including the category of HSIL+.

We can now look at a similar type of analysis for

determining specimen adequacy. This slide shows the results for truth adjudicated Satisfactory But Limited By slides. When comparing the two study arms, the AutoPap correctly detected 5,059 slides, which is greater than the 4,728 detected by Current Practice.

The equivalence test shows conclusively -- as illustrated down here at the bottom -- that the results show that the AutoPap-assisted practice is at least equivalent.

This slide shows the results for truth adjudicated Unsatisfactory slides. The AutoPap-assisted practice correctly diagnosed 137 slides and current practice detected 133.

As evidenced again by the p-value as illustrated here, the AutoPap-assisted practice is at least equivalent to current practice in the correct assessment of Unsatisfactory specimens.

While these data have demonstrated the improved sensitivity of the AutoPap-assisted practice, this increased sensitivity should not be achieved at the expense of specificity. Therefore, the issue of False Positive Rate becomes important.

Increased sensitivity with decreased specificity may be useful for the overall detection of abnormality, but it is hardly useful as an overall cost-benefit situation due

to the increased number of patient slides referred for cytopathologist review; or, for the increased costs in patient morbidity associated with any further diagnostic and follow up procedures.

This table shows the comparison of the false positive percentages between the current practice and the AutoPap-assisted practice. By false positive percentage, I mean the percentage of within normal limit slides that the cytotechnologists incorrectly classified as abnormal and referred to the cytopathologists.

You can see by the calculation at the bottom, that the AutoPap-assisted practice demonstrated a 16 percent improvement over current practice in reducing false positive cases. The difference between the false positive rates is statistically significant, as indicated by the p-value at the bottom.

As someone who runs a clinical laboratory, this difference is clinically meaningful for all of the reasons I have described earlier, including the reduction in cases reviewed by a cytopathologist, and the potential for reduction in unnecessary follow up tests for women with truly normal Pap smears.

What I have just presented represents the major findings of the study, which are based upon the pooled data

set. All performance claims and statistical results are appropriately based upon the pooled results. These results were observed across all clinical sites, and at each site, the AutoPap-assisted practice detected more abnormal slides than in the corresponding current practice study arm.

One question that may be raised is whether the AutoPap will miss abnormal slides in the AutoPap Review Only -- or No Review -- population. I would like to take a few minutes to address this question.

As an automated initial screener, the AutoPap is designed to assist cytotechnologists in identifying more disease, while enabling the laboratory to potentially process more slides and reduce the numbers of false positives.

In current practice and the AutoPap-assisted practice, a determination of no further review is either a cytotechnologist determining a slide is within normal limits, or the AutoPap classifying a slide as AutoPap Review Only, or No Review.

This slide shows false negative performance for the study truth results. The first column -- illustrated on the right side -- shows the AutoPap Review Only false negatives; the second column shows the final AutoPap-assisted practice false negatives -- that is this

column here. And this column actually includes the results from column from the first column.

Note that no HSIL cases, AIS, or cancer were missed by the device alone -- these are the results in column one. The important comparison, however, for the purposes of the study is the AutoPap-assisted practice false negatives -- column two -- versus the current practice false negatives, or column three.

In the AutoPap-assisted practice, only one HSIL and no cancers were missed, whereas in the current practice three HSILs and two cases of cancer were missed. In virtually all diagnostic categories, the AutoPap detected more disease.

Another issue to be considered is the contribution of initial screening, without including quality control rescreening performance. When the data for the study are analyzed, the AutoPap-assisted initial screening without quality control is at least equivalent to current practice. The addition of the quality control rescreening further improves the performance and added quality value for the laboratory.

It is also important to note that, in addition to this prospective study, which evaluated over 25,000 slides, NeoPath has already submitted to this Panel several studies

that evaluated very large numbers of slides from all Bethesda System categories, including approximately 600 HSIL and over 140 cancer slides.

These included the Current Archive Sensitivity Study and the Historical Sensitivity Study, both prospective and retrospective, and precision or instrument repeatability studies. Together, these studies represented data compiled on over 60,000 patients.

The AutoPap demonstrated significant sensitivity to all disease categories and showed 99.6 percent repeatability on multiple processing of the same slides on multiple AutoPaps. Finally, some questions have been raised, and even models developed, which presume that an automated device would not contribute to improved cytotechnologist performance.

This is because the models assume that the device does not provide information to the cytotechnologist. The data from our clinical trial demonstrate that the information provided by the AutoPap System does in fact improve the performance of the cytotechnologist.

Based on the data I have just presented, these are the claims we are making for the AutoPap System:

1. The laboratories detect significantly more LSIL+ Pap smears, including LSIL, HSIL, AIS and cancer, in

the AutoPap System-assisted practice, which is statistically superior when compared to current practice.

2. The laboratories detect significantly more LSIL alone Pap smears in the AutoPap System-assisted practice which is statistically superior when compared to current practice.

3. The laboratories detect significantly more ASCUS+ Pap smears -- again, ASCUS, AGUS, LSIL, HSIL, AIS and cancer -- in the AutoPap System-assisted practice, which is statistically superior when compared to current practice.

Finally, the laboratories reduced the number of Pap smears incorrectly classified as abnormal and thus, decreased the false positive rate by 16 percent in the AutoPap System-assisted practice compared to current practice.

In closing, I would like to state that the data presented are unequivocal. These data clearly show that the AutoPap safely and effectively improves the laboratory's ability to identify disease while potentially increasing efficiency.

Thank you very much. I would now like to introduce Dr. Tom Bonfiglio, who will discuss why we are confident in the conclusions of this study. Thank you very much.

DR. BONFIGLIO: Good morning. My name is Dr. Thomas Bonfiglio. I am a practicing cytopathologist and the director of a clinical cytology laboratory in Rochester, New York.

It is a pleasure and a distinction to be here this morning, and I hope to contribute to your deliberations.

NeoPath has paid my expenses to be here today, but I have no financial interest in NeoPath, or in the outcome of these deliberations.

I have personal experience with the AutoPap and very much want to see this application approved. Indeed, I think most other laboratory directors will also want the AutoPap as an immediately available option for enhancing their own programs.

This is primarily true because the AutoPap System is an important advancement in health care for women. This is not to say that the AutoPap is the perfect solution for enhancing the effectiveness of Pap smear screening.

For example, we realize that a very small number of the slides which are screened only by the AutoPap will be false negatives. But this is clearly outweighed by the device's benefits in reducing overall false negatives.

The AutoPap's benefits are clear. It helps a laboratory process accurately process more slides overall.

The AutoPap-assisted initial screening identifies at least as many slides as abnormal as the current practice, and AutoPap-assisted quality control rescreening identifies more abnormal slides than current practice.

In addition, the AutoPap helps improve the cytotechnologists' screening specificity. Not only does the AutoPap improve diagnostic performance, it is used with the slide preparation techniques employed by the vast majority of laboratories in the United States. Thus, the AutoPap provides a significant diagnostic improvement to women's health; fits within existing laboratory operations; and, meets our health care system's financial constraints by allowing a laboratory to process more slides, which helps compensate for the cost associated with the system.

All that I have just said regarding the device's benefits is based upon the unequivocal results of NeoPath's study, and upon the assumption that improving detection of Pap smear abnormalities is an indisputable clinical benefit.

It appears to me, then, that the issues before the Panel are limited. The questions are: whether the diagnostic improvements associated with the AutoPap are attributable to the device itself rather than to some artifact of the study; and, secondly, whether the AutoPap will continue to perform as well in commercial distribution.

My review of the clinical study indicates that the AutoPap, not some study artifact, caused the improvement in the screening process. My review also indicates that AutoPap's performance in the clinical trial is transferable to routine laboratory operations.

My conclusions are founded in the fundamental facts that have been discussed by the other speakers. NeoPath conducted a well designed prospective, masked, matched controlled study to evaluate the effectiveness of the AutoPap.

The study was conducted in an environment that ensured that AutoPap performance is transferable to other laboratories; and, the study delivered clinically significant and statistically significant results. These are the study design features that collectively ensured that any differences in outcomes are due to the AutoPap.

NeoPath's study compared the use of the AutoPap against the current standards for reviewing conventionally prepared Pap smears, including standard quality control rescreening.

The comparison was performed in the same laboratories, with the same personnel, over the same period of time, and with the same slides. The only apparent difference in the review process was the AutoPap.

NeoPath conducted its study in five large independent laboratories. Large independent labs screen the vast majority of smears in the United States. Using five laboratories minimized the effect of site-specific performance.

Using five laboratories -- large laboratories -- provided a larger number of cytotechnologists compared to using smaller laboratories for a longer period. This minimized the potential of the effect of the cytotechnologists' skills on study outcomes.

Using five laboratories helped ensure that performance under current practice would be representative of the industry. It also allowed for the use of more AutoPap devices, a total of eight. This helped minimize the theoretical possibility of superior or inferior performance of any one device.

During the study, the laboratories reviewed from 4,500 to 9,000 slides, each of which came from multiple pre-existing routine sources. No special care was taken in sample collection, in slide preparation or in selection. This ensured varying quality in the incoming slides, and ensured that the study was conducted in accordance with current practice.

The slides were first reviewed according to

current practice, and then reprocessed using the AutoPap. The cytotechnologists were masked to the current practice diagnosis during the AutoPap-assisted practice screening. Thus, the reviews were conducted on an independent basis.

The cytotechnologists were aware that a study was being conducted when they reviewed both the current practice and the AutoPap practice slides. Thus, the effect of clinical trial participation was equal in both arms.

Each laboratory participated in the study for about two and a half months. This was long enough to minimize any effect of performance created by the cytotechnologists' knowledge that they were taking part in a clinical trial.

The study included a truth determination plan. Whenever a slide was classified differently under the two arms, the slide was forwarded to an independent Panel of masked cytopathologists that determined the correct diagnosis. The Panel's determination was counted as truth during the analyses of outcomes. Thus, there was no classification bias in favor of the AutoPap.

More than 25,100 slides were reviewed. They came from women of different ages, different backgrounds, and different geographical areas. This gave NeoPath enough data for convincing statistical power, and for a dependable

analysis as to whether the AutoPap was indeed as good as, or statistically superior to, current practice.

To summarize, the study was performed in enough laboratories, involved enough sources of slide preparation, involved enough cytotechnologists, involved enough devices, occurred during a long enough period of time, and utilized enough techniques to prevent bias, and included enough slides to ensure that the difference in results did indeed occur as a result of the AutoPap.

The next question is whether other laboratories will benefit from the use of this device. I think we can expect the AutoPap to provide the same level of improvement in other laboratories, unless these laboratories differ in their current practices, or use the AutoPap in a different way.

Large independent laboratories, like the sites that participated in the study, screen, again, the vast majority of Pap smears in the United States. NeoPath's control arm represented current practice. Its sites represented how most laboratories process Pap smears.

NeoPath did not compare the AutoPap against the use of liquid-based prepared slides because this represents less than 10 percent of slide preparation methods currently. It did not compare AutoPap against the quality control

version of the AutoPap because this is not considered the standard, either. NeoPath selected its control arm based on its prominence; that is, the prominence of the control arm within the industry.

I also believe other laboratories will use the AutoPap as it was used in the clinical trial. I think the important variables affecting this question are the quality of a laboratory's slides, the skills of the cytotechnologists, and its workload.

By workload, I mean the time each cytotechnologist has available to screen slides. We all know that workload can affect the quality of reviews. In other laboratories, cytotechnologists will review as many slides under the AutoPap-assisted practice as are reviewed under the current practice.

Here again, the AutoPap's performance in the clinical trial will be transferable. Slides for both the current practice and the AutoPap-assisted arms of the study were reviewed by cytotechnologists as part of their normal workload. The workloads in the laboratories were not reduced, nor were cytotechnologists allowed more time to review the study slides. All this was done to simulate laboratory use so that results would be transferable to use in commercial distribution.

I want to remind you of the study's statistically significant and clinically important results. False negatives were reduced from 21 percent to 14 percent, an improvement of 33 percent. More importantly, false negatives were reduced for the clinically important disease categories.

For LSIL, the AutoPap reduced false negatives from 16 percent to 9 percent, a calculated improvement of 44 percent.

For low grade lesions, high grade lesions, and cancers, false negatives were reduced from 14 percent to 8 percent, an improvement of 43 percent.

The AutoPap System also improved specificity as well. The false positive rate decreased by over 16 percent, and this as you heard, is statistically significant.

In closing, I would like to reiterate and concur with Ms. Norton's remarks from the beginning of our presentation. The AutoPap should be approved without delay because: NeoPath has conducted a prospective, masked, matched controlled study, consistent with FDA, Panel, and Intersociety Working Group recommendations.

The NeoPath study provides clinically and statistically significant and unequivocal results. The NeoPath study was conducted in laboratories that operate

under current practice standards. For this reason, other laboratories will experience the same enhanced performance with AutoPap.

The AutoPap is designed for use with the standard slide preparation techniques that the vast majority of laboratories use in the United States. Thus, it is a technological enhancement that will have practical application throughout the country.

And finally, and perhaps most importantly, the AutoPap will assist laboratories in reducing false negatives while also reducing false positives. It will help save lives. It will significantly advance health care for women.

This concludes our prepared remarks. Thank you for your attention. NeoPath will be happy to answer any questions. Please direct these to Ms. Norton, who will ensure that the most knowledgeable person is able to respond.

DR. FRANCIS: Thank you, Dr. Bonfiglio and the other members of the NeoPath presentation team. We are running a little bit over time, but I do want to give the Panel just one opportunity, if they have any issues they would like to raise now instead of holding them over to the discussions this afternoon, please so indicate.

PANELIST: [Question away from microphone.]

DR. FRANCIS: No, the sponsors will in attendance this afternoon, and you will have an opportunity to raise your questions then, but if there is just something that you really think you ought to bring up now, rather than holding over lunch. Yes? Would you state your name for the record?

DR. DAVIDSON: I ask this question because I am not sure that people will be here this afternoon. Are there any significant differences between the criteria established with this Intersociety Working Group versus ASCP?

DR. FRANCIS: Is there no one in the audience that can respond to that question? Dr. Stanley.

DR. STANLEY: You will forgive me, I have a compelling invitation to lunch. Of course, this document was published in essentially unpublished form in the ASCP Journal, as it was in a number of other places, as certainly it should have been.

It was then editorialized about, and I think a number of issues were underlined or brought up, but the ASCP wishes to go on record as supporting that document as it is published. Have I answered the question that was asked, or have I answered some other question?

DR. DAVIDSON: I do not think so. Are there any significant differences between the ASCP criteria and this Intersociety Working Group criteria?

DR. STANLEY: The ASCP wishes to go on record as supporting the criteria as published by the Intersociety Working Group.

DR. FRANCIS: Dr. Davey.

DR. DAVEY: Yes, actually, I wanted to bring up one thing that is more from the public comments. I do not know if we can clarify this, but I was -- the Trylon Corporation statement that you read, I was at this -- I just wanted to say that I do not agree -- I was at the July 1997 Panel meeting, and I do not think that the comments made by Dr. Lonsky are really a true representation of the conclusions.

If one of the FDA staff that was there wants to look this up or comment on it, I just wanted to go on record as saying that I do not think that his comments were a fair representation of what the gold standard should be.

DR. FRANCIS: So noted.

DR. BIRDSONG: I was not at the Panel meeting, but I would like to second Dr. Davey's comments. This is the first time I had seen comments like that. I do not find them representative of the field.

RABINOWITZ: Dr. Max Rabinowitz. I am a medical officer in the Division of Clinical Laboratory Devices, and also, am involved in the review of in vivo devices for

cervical diagnosis. So, I was at that Panel.

I think the thinking of asking for a biopsy confirmation for new in vivo technology was to validate the new in vivo technology, visualizing the cervix with different optical, electro-optical, and other means.

Whereas, what we are dealing with today is looking at the very same technology with computer enhancement. So, I think our in vitro diagnostic devices are not to validate cytology, but rather -- so, we felt it was proper to have cytology -- adjudicated cytology -- as the reference method.

Whereas, looking at new, novel devices for making diagnoses required a reference method to adjudicate between the in vivo and the cytology methodology. I hope that is clear.

DR. FRANCIS: Thank you.

DR. DAVEY: It is not the same thing and I think that we [comment off microphone].

DR. FRANCIS: If there are no other questions or comments at this time, I will adjourn the meeting for lunch and we need to reconvene here at 1:30 sharp.

[Whereupon, at 12:40 p.m., a recess was taken until 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. FRANCIS: I would like to welcome everybody back to the second half of today's Panel deliberations, and in the first part of this session, we are going to hear from two members of FDA staff. After that, I will open the floor to the Panel to ask questions of our sponsors, whose presentations we heard just before lunch, and of the FDA staff.

I would like first to ask Michelle Stuart, Scientific Reviewer, Immunology and Pathology Branch, Division of Clinical Laboratory Devices, Office of Device Evaluation, to address the Panel.

Agenda Item: FDA Presentation

MS. STUART: Good afternoon. Today we are here to review the AutoPap System submission from NeoPath. If approved, this would represent the first automated cervical cytology device intended for use as a primary Pap smear screener.

In 1995, NeoPath originally submitted the AutoPap 300 QC System, which was approved by the FDA for rescreening of all within normal limits or negative slides on September 29, 1995.

The firm submitted the AutoPap System submission for use as a primary screener in 1996; however, this device

submission was considered not approvable by the Hematology and Pathology Devices Panel at the September 27, 1996 meeting.

After consultation with the FDA and some of the Panel members, NeoPath submitted on August 28, 1997 the prospective intended use clinical study data that we have asked you to review and that is the focus of our meeting today.

As with all previous submissions from NeoPath, the AutoPap System is intended for use on slides that are conventionally prepared.

Now, take a look at the Indications for Use. The AutoPap System is an automated cervical cytology screening device intended for use in primary screening and quality control rescreening of Papanicolaou or Pap smear slides.

The device is to be used only on conventionally prepared slides and is intended to detect slides with evidence of squamous carcinoma and adenocarcinoma and their usual precursor conditions. These abnormalities fall within the following diagnostic categories of The Bethesda System. The categories are: AGUS, ASCUS, LSIL, HSIL, and Carcinoma.

There were some laboratory exclusions when the AutoPap System trials were going on, and some of these were slide limitations, and because of these limitations, they

eliminated some of the slides from the prospective intended use clinical trial.

Of the slide exclusions, the reasons were high risk, AutoPap limitations, laboratory exclusions, and incorrectly processed. Some of the laboratory exclusions contained broken or cracked coverslips, plastic coverslips, and cases with multiple slides; or specimens that were collected only with the endocervical brush; missing slides or source documents; or the slide was not a Pap smear. So, consequently, there was a total of 5,336 slides out of the original 31,507 slides submitted for evaluation that were excluded from this study.

This device and its predecessors are not intended for use on slides of high risk cases. The reasons some of the slides were called high risk are shown as abnormal Pap, abnormal bleeding, biopsy, cancer, chemotherapy, colposcopy, radiation, HPV, high risk not otherwise specified, directed screen, other history comments, other directed comments, and in addition, AIDS patients were also put into the high risk category.

As you can see on this list, the most frequent reasons for high risk slides were the top three; abnormal Pap, abnormal bleeding, and biopsy.

Other limitations were due to the device being

capable of analyzing the slides, and these included bubbles over more than 5 percent of the area; marks, dust, prints or adhesive; writing over the barcode; labels over slide edges; labels overlapping the coverslip; the slide or the coverslip being of the incorrect size.

Now let's get into the current practice protocol. Out of the original 31,507 slides, I mentioned before that 5,336 were excluded. In this above protocol, 26,171 slides were the total number, then, that were entered into the study. Out of this total number of 26,171, there were 963 that could not be processed by the device, meaning that the final total ended up being 25,208 that qualified for AutoPap screening.

In the AutoPap Assisted Practice Protocol, 25,208 slides met the criteria for AutoPap System processing; 5,109 of these slides were placed in the No Review category by the AP System, and that is approximately 20.26 percent.

There were some laboratory exclusions that I had mentioned before, and consequently, I will not repeat them, but they are still listed on the board if you want to see them again.

The No Review category for the AutoPap System contained the following disease categories: Within normal limit slides, there were 5,011; unsatisfactories, for unsats

41; ASCUS 41; AGUS 2; LSIL 14; giving that total of 5,109 in the No Review zone. And we would like to mention that there were no HSIL, AIS, or Cancer cases included in this No Review zone in the study.

This slide is just to show you that when you look over the five clinical study sites at the No Review criteria, the lowest one was 16.13, and the highest one 24.46. When you look at the QC Review category, then you see that the lowest one is 12.79, the highest one is 21.29.

At this point, I would like to introduce Judy Chen, who is our statistician for the review, and she is going to give a statistical summary, and then I will come back to conclude the FDA presentation.

MS. CHEN: Now we will take a look at the data of the prospective intended study submitted by the sponsor. Before we look into the data, we need some definitions.

We have seen in a similar slide in the previous presentation that the AutoPap process and the conventional process divided slides into four categories, and they can be grouped as seen in the table: positive-positive; negative-negative -- these are the concordant slides. And also, the discordant slides: positive-negative; positive by AutoPap process, but negative by the conventional process; and vice versa.

An important measurement is the positive-negative over the total discordant slides. Also, later on we will talk about the two -- for example, ASCUS/AGUS. And the two in here are defined as concordant diagnosis, plus the EDP diagnosis of the discordant cases, it is not biopsy results.

As you can see, for the ASCUS/AGUS cases, in each of the five laboratories, the proportions are numerically all larger than 50 percent, but statistically, the two-sided 95 percent confidence interval also indicated that the lower three of the five laboratories actually are higher, or just make 50 percent; that means, even by laboratory, these laboratories showed a statistically significant difference, favoring AutoPap.

Since there is no statistical difference in these five proportions, it is reasonable to pool the result. As you see in the last row, the pooled result indicates a statistically significant difference, favoring AutoPap, for which the proportion will always be higher than 50 percent - - the lower band is 54 percent.

For LSIL, the pattern is similar; the only difference is that by laboratory, none of the laboratories by itself shows a statistically significant difference, but since there is no difference among the five laboratories, the pooled result does show a statistically significant

advantage for the AutoPap process, over the conventional process.

Here, for HSIL, AIS and Cancer, since the numbers are so few by laboratory, you really cannot say anything. The pooled result, numerically, favored AutoPap, with the proportion five over seven, and the confidence interval is .29 to -- and .96. This did not show statistical significance.

The last slide will show you the relative sensitivity by disease category. The reason I put the relative there is the gold standard really here is not biopsy result. And you can see the AutoPap outperforms the conventional process for AGUS/ASCUS, and also for LSIL. For HSIL, AIS, and Cancer pooled together, there is no statistically significant difference.

MS. STUART: Before I conclude the FDA presentation with the questions that we have for Panel deliberation, I would like to thank our Division Director, Dr. Gutman, our medical officer Dr. Rabinowitz, my Branch Chief, Dr. Maxim, and all of the other review members on my team -- Mary Anderson, Larry Brenza, David Brown, Judy Chen, and Louise McGruder. Thank you.

Now we will go on to the questions for Panel deliberation. The first question we would like you to

consider is:

Do the data presented in this PMA support the manufacturer's intended use of the AutoPap System?

The second question that I would like to get input from you for is, are the claims being suggested by the manufacturer for their device's performance compared to manual screening supported by the data available? If so, how should these claims be presented in the labeling, and what, if any, limitations should be applied to this data presentation?

Our last question for Panel deliberation is, in the intended use study, the AutoPap System provided a ranked report to the screening cytotechnologist for all reviewed slides. Will having knowledge of the ranking affect the cytotechnologist's vigilance?

That is the end of the FDA presentation. Thank you.

Agenda Item: Open Committee Discussion

DR. FRANCIS: Thank you very much, Ms. Stuart. I would like to open all the presentations now for questions and discussion and perhaps we could -- if there are any questions to be asked directly of the FDA staff members or the sponsors here today -- we could address those issues first before we address specifically the questions raised by

the FDA. Would anybody like to kick off? If you would state your name for the record, please.

DR. DAVIDSON: Yes. Davidson. I wonder if someone could provide a perspective -- I am referring to the discussion of sensitivity by Dr. Garner from Canada in the public session earlier -- indicating that the sensitivity of the screen was not improved unless the cytotechnician's vigilance was improved.

I wonder, in regards to that conversation, how he or someone would interpret the current data that is being presented by AutoPap.

DR. FRANCIS: Would anybody like to respond to that question?

DR. BONFIGLIO: I would be happy to try and respond to that. The information that he discussed was similar to the model that was published by the Intersociety Working Group, of which I am a member. But the assumption in that model was that the automated screening device would not provide or produce any increased sensitivity on the part of the cytotechnologist.

I think the data you have seen today suggests -- quite strongly in fact I think demonstrate -- that the data does not fit that model, that information provided to the cytotechnologist through the use of the AutoPap screening

device does indeed increase the sensitivity of the cytotechnologist.

I think we can say that, apparently it is increasing the vigilance of the cytotechnologist, at least according to the data we are looking at.

DR. DAVIDSON: In the -- I think it is the Palomar Study -- in this question of what impact would ranked results have on the quality control -- and I understand that the higher classes would include the higher numbers of abnormal findings.

If you were to look at that table, in the higher class, 1 out of 4 from the 1 to 20 were abnormal, if you eliminate the unsatisfactories. And in the second, 20 to 40 rank, 1 out of 6 was abnormal. In the 40 to 60, 1 out of 4 was abnormal. And in the 60 to 80, 1 out of 3 was abnormal. In the bottom class, there was 1 out of 10.

It appears from the discussion that this would be a progressive decrease in abnormalities, but when you look at the specific ratios, it does not increase that way.

MS. NORTON: I am Mary Norton. I would like to comment. That was a preclinical assessment to determine the feasibility of the ranking concept. The number of slides in that study were not sufficient to have a statistical significance, as in the current intended use study. So, the

preclinical study was conducted to give us an estimation of the value of the ranking and a general concept of what we might expect in the larger prospective intended use study, where we could study this on a larger number of cases. However, the trend, as you indicated, does reflect the correlation with the lower ranking slides having the higher probability of being abnormal.

DR. DAVIDSON: Do you have other data, then, that supports this hypothesis, that, you know, the ranking and --

MS. NORTON: In the intended use study, we in fact looked at outcomes by rank. We have a few slides prepared to show you -- I believe they are also in Volume 9 of the application -- that shows by the disease categories tested, that at the lower ranking, or for the higher probability of abnormality, that we gained a larger number of cases as compared to current practice. Would you care to see the slides on that?

DR. FRANCIS: Would you like to see those slides, Dr. Davidson?

DR. DAVIDSON: Well, if it does not take up too much time.

DR. FRANCIS: We have the time.

MS. NORTON: If you will just give us a few minutes here to get them.

DR. FELIX: Can we entertain a brief question in the meantime?

DR. FRANCIS: Sure. Shoot.

DR. FELIX: I would like to understand a little bit better. In your analysis of the false positives, you showed that those slides processed by the AutoPap had a slightly smaller degree of false positives, and I think it was 4.-something versus 5.-something. Now, how was that determined? I am not quite sure how that is determined; I missed it in my review of the data.

MS. NORTON: I would like to have Dr. Richard Chiacchierini comment on that question.

DR. CHIACCHIERINI: I am Dr. Richard Chiacchierini. I am the Vice President for Statistical Services of C.L. Macintosh and I am a consultant to the company. I have no financial interest in the company, other than my fee-for-service consulting agreement.

The false positive rates are based on a table, very much like you saw in the tables that the primary clinician presented for this particular approach. Could I have the false positive table, please, Tim?

Okay, well, basically, there were 23,556 negative slides. Of those, 1,113 were called false positive --

DR. FELIX: Can you start all over again? Sorry.

DR. CHIACCHIERINI: I will. I will. There were 23,556 negative slides. Of those, there were 1,113 called false positive by AutoPap; there were 1,323 called false positive by current practice.

The two cells of the table in which we are particularly interested are the discordant cells; and the discordant cells in the lower left-hand corner of the table would be 1,212 that were called within normal limits -- I am sorry -- that were called false positive by the current practice, which were truly within normal limits or unsatisfactory. And there were 1,002 slides that were called false positive by AutoPap, and were called within normal limits or unsatisfactory, that were truly within normal limits or satisfactory.

The test is a test based on the discordant pairs; that is, the sum of 1,212 plus 1,002. And if you do that test with an exact binomial or a chi-square against a null hypothesis -- there it is, right there -- I am sorry -- you are actually comparing 1,212 versus 1,002, because 1,113 can only be different from 1,323, if those two numbers differ.

The Macnamar's(?) Test or the exact binomial test with an observed proportion of P equal to .5, gives you a highly statistical difference; the chi-square is over 19.

DR. FELIX: Thank you, that's great. I am

satisfied.

MS. NORTON: To address Dr. Davidson's questions, we now have those slides prepared, if you would like to see them.

DR. FRANCIS: Go ahead.

MS. NORTON: This is the first slide, which shows the performance on the first tested category ASCUS+, and what you see in the left column are by rank with the rank of 1, or 0-20 representing a slide with the highest probability of abnormality, and 5 representing the lowest probability for being abnormal. And again, as you see indicated below, the column designated as No Review represents the slides that would be reviewed by the AutoPap only.

The two columns adjacent to that, indicate the number of cases gained by the AutoPap-assisted practice, versus the current practice for all of the ranks indicated on the left. So, if I were to -- yes?

DR. DAVIDSON: But I -- my question had to do with also knowing what the denominator was, the relative risk. See, what drew my attention, for example, in the 60 to 80 group, in the Palomar Study, although it is a small number of slides, one out three was abnormal. Which means that you -- though it may be a small number, that is not a category if that were to hold, that could you differentially discard

that on a quality basis, and look towards the upper part of that column.

MS. NORTON: If I could just take one moment to confer. I hope I am responding correctly to the question, but this represents a percentage of ranking, so approximately 20 percent in each of the rankings of slides is associated in those columns. Dr. Bonfiglio, would you like to comment?

DR. BONFIGLIO: I think the point is, these are quintiles, so there is 20 percent of the total slides in each of the five blocks, from 1 to 5. So that there is the same total number of slides in each category, the highest quintile having a higher percentage of abnormal slides.

In the AP-assisted arm, the cytotechnologist is looking at all of those except for the group that says, No Review. So, nothing is discarded. Those slides are all reviewed, but the highest quintile, as depicted here, contains, obviously, the largest number of the abnormal slides.

DR. DAVIDSON: I thought that, in the normal category, that this ranking was in an effort to direct the attention to the cytopathologist as to which category you are more likely to find abnormal slides in.

MS. NORTON: That is correct.

DR. DAVIDSON: And so that if this were to influence the cytotechnologists, they would look more at the slides near the top of the column, and less at the bottom, is that correct?

MS. NORTON: We would not necessarily know that that was the case, but in fact, the one would be associated with the highest probability of being abnormal, and five, the lowest probability of being abnormal, but still classified as review by the AutoPap System.

DR. DAVIDSON: Yes, but see, my question has to do with, though it may be 42 in the lower category, from a percentage standpoint, those may be more abnormal -- I am just looking at the table -- you know what I am referring to.

MS. NORTON: Yes.

DR. FELIX: So there is about 5,000 cases in each ranking, about -- for the quintile? You start out with about 25,000 cases --

DR. BONFIGLIO: No. No, that is not true.

MS. NORTON: It -- right. In the top 75% of the cases, the 75% of the cases designated as review, those were the pick-ups at every quintile.

DR. BONFIGLIO: So, out of the 25,000 slides, 25 percent are in No Review, the others are divided in those 5

quintiles.

MS. NORTON: But they may not necessarily be evenly -- about 4,000.

MS. ROSENTHAL: So then, would you suggest for the 15 percent quality control review you would --

DR. FRANCIS: Could I ask the Panel members to please use the microphone?

MS. ROSENTHAL: So, then you would be asking for quality control review on the top quintile, is that it, in -
-

MS. NORTON: That is correct. That the quality control rescreening slides would most likely occur in the top quintile, or possibly into the second quintile, if in fact the cytotech had determined a number of those slides to be abnormal, and so they were removed from consideration, and the next slides that were designated as normal would be selected. So, that is correct. In the top quintile, possibly some into the second.

DR. FRANCIS: Dr. Davey.

DR. DAVEY: Diane Davey. Could I just clarify; the same exact slides are in both the AP and the CP -- I mean, the slide ranking is dictated by the AutoPap, and it is just -- I mean, because there is really no -- there is really no scoring for the CP, and so it is just -- it is

exactly the same slides and -- okay. I just wanted to make sure.

MS. NORTON: That is correct.

DR. FRANCIS: Any other points on this particular issue? Dr. Davey.

DR. DAVEY: Actually, if we are talking about it. Would you recommend -- if this is used -- I mean, I guess I am a little concerned about the vigilance of technologists when it gets out in the field, if they will not be as careful if they have a low score slide, at least with the initial screening. I would think you would want to use it for the rescreen -- you would have to use it for the rescreening, but would you recommend reporting the score on the final report?

Also, what if -- how would you report a slide that did not get any review, in terms of reporting it clinically? Would the patient know, in other words, and the clinician know that it just went through the AutoPap?

MS. NORTON: Would you like to comment on that, Dr. Wilbur?

DR. WILBUR: Well, there are any number of ways that one can do this, and I do not believe at this point that NeoPath has made any particular recommendations on how that would be done; certainly, as a laboratory director, I

would be responsible -- as any laboratory director would be responsible -- to ensure the quality that goes through my laboratory, and should I --

If you are asking me, would I report whether AutoPap only was designated on the report versus manual screening, or even manual QC screening? I guess I would ask the question back; do we report that a slide has been QC'd, or do we not report that a slide has been QC'd on the final report? If you do, then obviously you have considered that that is important for the clinician to know. If you do not, you consider that it is irrelevant, because it is basically an internal laboratory policy.

The question is, do you consider that it is important that the clinician know that information, and I could -- I am sure that if I sat here, I could argue it both ways. And perhaps you could, as well. Dr. Bonfiglio?

DR. BONFIGLIO: The way I look at this, based on the study, I think it might be important to report that these slides were only reviewed by manual rescreening, because the data would suggest that it is better to have it reviewed by the AutoPap.

DR. FRANCIS: Dr. Davey? Anybody else care to comment?

DR. ALLEN: This is Dr. Allen. Could you, once

again, recapitulate the importance or need to group the slides, if this device is utilized in general practice? Or would you still group them in quintiles? Or, is this just for research purposes?

MS. NORTON: It would be optional for the quintile designation to be provided. Currently, and as was conducted during the study, they received the absolute ranking of the slide within the population of slides processed at that time, but it would be optional.

DR. BIRDSONG: I guess I am having a little trouble with you saying it would be optional. It seems to me from the data presented that it is a pretty key part of the protocol and you really could not have anything optional.

DR. BONFIGLIO: But, I think she is referring to quintile, but I think the ranking is important because I think that is what --

DR. BIRDSONG: The quintile is what I am referring to here.

DR. BONFIGLIO: Well. They are both giving you the same type of information. If it is high ranked, it is -- I think either information is relevant to report.

DR. FRANCIS: Other comments? Dr. Renshaw?

DR. RENSHAW: I just had two real questions. The

first is a procedural one. I was hoping you could walk me through a slide through your study. I just wanted to make sure I understood how it actually went.

A slide arrives in your lab. It gets assessed(?), everything the same as usual. It is first read in the normal way -- is that correct? In the routine manual screening?

MS. NORTON: That is correct. I can respond and then Dr. Wilbur would like to add comments.

The slides came into the laboratory, as you correctly pointed out, via the normal assessmenting(?) processes. They were evaluated against the protocol inclusion and exclusion criteria, and were either included or excluded, based upon those issues.

They then were processed in the standard current practice of the laboratories, as they would have been screened without the AutoPap in place. They received no additional information, obviously, only that which was provided by the laboratories or the clinical information on the patients. And a cytotechnologist screened the slide, just as they would in normal practice.

The difference is that at the end of that process, those same slides were then run on the AutoPap instrument -- 100 percent of them processed on the AutoPap -- and 25

percent of them were designated as AutoPap review only. And the 75 percent approximate slides remaining were sent to a screening cytotechnologist for their determination of the slide.

Based upon that initial screening, as on the current practice side, some were subjected to the additional quality control rescreening; in this case, the quality control rescreening was dictated by what the AutoPap instrument selected for quality control rescreening.

If the slide which was within normal limits was not selected for QC rescreening, those results were recorded. Of course, all abnormal were referred to a pathologist for our EDP process for truth determination, if it was found that the cytotechnologist screening this slide on the AutoPap-assisted arm had a different determination than that of the current practice screening cytotechnologist. So, for either diagnostic discrepancies or adequacy discrepancies, those were referred to the external Panel for adjudication.

DR. RENSHAW: Right. But there was no additional rescreening of the slides anywhere in between those steps.

MS. NORTON: There was not.

DR. RENSHAW: So, if I am correct, the regular route was always first. The routine manual thing was first.

And the AutoPap was always second. Is that correct?

MS. NORTON: In terms of the order?

DR. RENSHAW: Yes.

MS. NORTON: That is correct. They would be processed on the current practice and then screened by the AutoPap. They were two independent processes.

Now, it would be the case that, on some days in any given laboratory, different slides were at different points in the process, as you might expect. This was a continuous operation that ran every day, but that is the case.

DR. RENSHAW: And all the dots from the routine screening were removed before it was put on AutoPap.

MS. NORTON: That is correct.

DR. RENSHAW: So when the slides were sent to the EDP, the only dots that were on it were those from the AutoPap review, or did I miss a spot.

MS. NORTON: -- no, that is not true. We actually saved the dots from the current practice and we dotted the slides with both sets of dots so as not to bias them in their determinations.

DR. DAVEY: The dots were the same?

MS. NORTON: I am sorry?

DR. DAVEY: I am sorry. Were the dots -- did the

dots look identical?

MS. NORTON: I could not comment on that. I did not read the slides.

DR. WILBUR: I think there are two points that perhaps we are missing. One is that, if I am understanding perhaps what your concern is, is whether or not the cytotechs that reviewed in the current practice actually knew they were reviewing study slides, or whether they were reviewing part of their normal workload.

Well, they did know that, so I think that is an important piece of information. A slide was assessed into the study so that both the current practice cytotechnologist as well as the AutoPap-assisted cytotechnologist knew that they were in a study slide situation.

The second part of your question is, the EDP pathologist did not know which sets of dots were coming from the CP side or AP side.

DR. FELIX: And you ensured that those dots were placed in the exact position via what methodology?

Coordinates --

MS. NORTON: We actually xeroxed the slides with the dots on them and the transparency was overlaid on top of the slide and the cytotechnologist replaced the dots as they

were on the original screening.

DR. FRANCIS: Dr. Williams?

DR. WILLIAMS: I had another procedural question. The slides that were excluded, were they looked at by either the cytotechnologist or the AutoPap at all? Or, once they were excluded out of the study, they were more or less forgotten?

MS. NORTON: They were excluded from the study and they were not analyzed on the AutoPap-assisted practice study arm. They, of course, would have followed and been screened during the current practice, so that results could have been reported to the patient, but they were not included in the AutoPap-assisted arm of the study.

DR. WILLIAMS: Okay. Were there any studies, or was there any data collection on those slides that were excluded?

MS. NORTON: We showed earlier the data for the slides that were excluded from the study, the 5,336 slides that were referred to earlier by the reason for their exclusion, and they were excluded based on the predefined protocol exclusion and inclusion criteria, as established in the protocol and by the laboratory itself.

DR. FRANCIS: Dr. Davey?

DR. DAVEY: I wanted to, I guess, talk a little

bit about the high risk category, and a few concerns. I can understand the philosophy of excluding them, in that, if you have a high risk patient, you are going to want to screen them and possibly rescreen them if they come to your lab, and if they go into the No Review pile, you are left sort of with a problem.

My problem is, is that the high risk criteria are so variable, and a lot of times we do not know the clinical history of a patient, so you know, you can come to a number of situations where you would have a patient come in with very little history, and it would go into the AutoPap and then you find out later with some additional history, the patient was high risk.

Then do you have to go back and relook at the slide -- or, what if that patient is found to have cervical cancer and someone said, well, but you should have known this patient was high risk? It just -- it causes a number of problems, I think, with implementation, and I guess that -- I was sort of wondering if you had any data now.

It seems to me that probably some of these patients that had abnormalities probably were indeed high risk patients that you did not know were high risk. However, you know, when labs are putting this into practice, it would sort of promote -- cause problems, especially if

you have a high risk population, of knowing who to put in and what to rely on; how much clinical history do you need, and you know, HCFA has come up with high risk criteria for who should get -- who gets covered for screening Pap smears, and all that, and there are all those things.

I just wanted, you know, some comments. It would have been helpful, I think, if we would have at least had the high risk slides run through just to know what happened. I mean, you can exclude them later, and so I was curious if any of that data was available.

PANELIST: That was my question.

DR. FRANCIS: Good question.

MS. NORTON: Would you like the sponsor to comment on that?

DR. DAVEY: Yes, and maybe the FDA, I -- I guess, you know, those are just concerns about what happens if you have later a patient that you find is high risk that went through this? What -- you know, how is that --

DR. GUTMAN: We will let the sponsor take that.

DR. FRANCIS: Sponsor commenting?

MS. NORTON: Dr. Wilbur, would you please comment?

DR. WILBUR: Well, I actually think -- I think Dr. Davey actually understands the principle, as she states it, that high risks are a difficult problem, but each site in

the study has its own definition of high risk. And therefore, based on the current practice in that laboratory, they set a typical -- or a different -- standard from their current practice.

A high risk slide may be screened manually twice. It may always go for QC. It may always be manually screened, in a protocol that even includes AutoPap, even though such slides are categorically excluded in the AutoPap labeling.

So therefore, there is a separate standard of care for high risk cases in each laboratory. And again, that is determined by the laboratory.

To get back to your point about what if a -- you know, a lot of these cases clearly have high risk criteria that we just do not know about. Well, in fact, the intended use study was large enough so that, undoubtedly, some of those cases that were detected were included, and the data therefore -- if you look at it that way -- does include some hidden, if you will, high risk slides.

That is exactly what we would expect in current practice. There would be hidden slides that were high risk, and in the study, there were hidden slides that were high risk. So, in fact, we have already addressed that issue, just by the way the protocol was developed. Is that a fair

answer?

DR. DAVEY: Well, yes, okay. So, then what if a laboratory then finds out later that the patient is high risk, then what would happen?

DR. WILBUR: Well, again, I think that would be up to -- I mean, you are talking about an historical way to define a high risk slide. You do not define a high risk slide by anything other than history; history of an abnormal pap; abnormal clinical history of some kind or another.

It is not something intrinsic on the slide that defines that slide as high risk, so therefore, that information is not available to the AutoPap per se, it is only available to the person that is reading the slide, and therefore, it is something that takes that slide out of current practice.

DR. DAVEY: Yes, I guess what I am getting at is the circumstance comes in; you get the patient -- you know, the typical thing is, you get the patient's name. They may put the LNP(?) on there, you know they have all the correct identifying information that is required to accept the slide in the laboratory. But then, you put it through.

It goes in the AutoPap No Review category, and then two weeks later, the clinician calls you and says, oh, by the way, we forgot to put the history on this slide.

Yes. And then do you pull it out again, and manually screen it, or what do you do, I guess?

DR. BONFIGLIO: Well, I think I might probably do the same thing if that happened now and a slide went through as negative, and the clinician called me and said, do you know that patient was high risk? I would probably go back and look at it again, if the physician was that concerned about it to call me and tell me that.

I think the important point is, though, that, you know, the high risk cases, the disease in the high risk patients is the same as the disease in the, quote, low risk patients, and the appearance of the cells is the same. So, I think we can logically assume that the instrument will be able to identify high risk patients who slipped in there. I mean, we identified them in all these other cases, there are clearly some high risk patients in this group.

I am not concerned about it in regard to safety. I think each laboratory is going to have to make up their own -- criteria for high risk, and how to handle those high risk slides.

DR. FRANCIS: Ms. Rosenthal, did you want to comment?

MS. ROSENTHAL: Yes. I have a little problem with that, because I think that the data that you are presenting

-- I assume you are presenting it to be used as you used -- to have this used as you used it. And there is a tremendous assumption that the clinician is going to communicate to the laboratory that it is a high risk patient and that the laboratory is going to understand what to do with a high risk patient.

It certainly would have been helpful to have put those high risks through the AutoPap system, just to indicate that, in fact -- or, to show us exactly what would have happened with high risk patients.

I am a little uncomfortable. I think that, you know, if I were, say, Carol Ann Armenti, I would want to know that this has been tested to the nth degree, and I also would think that I would want my gynecologist to receive a report that indicated that in the quality control rescreen, this had a high percentage -- this was ranked high, even if the cytologist and pathologist decided that it was within normal limits.

DR. FRANCIS: Dr. Felix, do you want to say something?

DR. FELIX: Yes, I have a question regarding the same issue. I understand the rationale that you have explained to us about the note eval- -- excluding the high risk patient. What I am not quite as clear in my

understanding is, how it is going to affect the indication for usage, or your indication? Because there are various functions of the instrument, one of which is already approved, which is the QC.

In laboratory practice, if the FDA approval is granted for primary screening at 25 percent -- and I am posing a theoretical question -- you put 100 percent of the slides through, whether they are high risk or not. Because you are also doing this for QC, and you have to put 100 percent of the normals -- or the, within normal limits -- through.

Let's say that you have run all of your slides through. What would be your -- at what point would you exclude slides, once you get results of within normal limits, and then you find out that the patient is high risk? At what point in the process do you exclude that slide from analysis?

MS. NORTON: The study was conducted so that in fact, the high risk cases were excluded prior to any running on the AutoPap, so the comment referring to 100 percent of the slides processed on the instrument meant that those were the slides after the high risk slides had been removed from the population.

In the current approved use of the quality control

system, the directed rescreen of high risk population, also remains as a limitation on the product. So, for the quality control system, those slides are excluded first and then the remainder of the negative slides are processed.

DR. FELIX: So, it is my -- my current understanding now after your response that none of those slides should go through the AutoPap.

MS. NORTON: As we conducted the study, and as we would expect the labeling to read, that is the case.

DR. FELIX: And that is going to be the labeling as it reads.

MS. NORTON: The case -- that is right. It has a limitation for high risk cases.

DR. FRANCIS: Did any of the slides subsequently turn out to be high risk? As Dr. Davey was saying, did you get information later on any of those slides which would have you reclassify those as high risk?

DR. BONFIGLIO: I am sorry. Can you restate the question?

MS. NORTON: Can you restate that?

DR. FRANCIS: Your high risk slides were removed prior to the study, but were any slides subsequently reclassified as high risk, after they had been through the AutoPap, to give you any small body of data on high risk

slides?

MS. NORTON: They were all removed from the study, so we did not collect information.

DR. FRANCIS: So, no information came after that time?

MS. NORTON: That is correct. I would like to make one comment, though. For the quality control system, we did show that the AutoPap had no particular bias against high risk processing of slides, and in fact, you know, we would be open to high risk slides being run on the instrument, with the exception that, even though they had been processed on the AutoPap, that they would then have to be subject to whatever special procedures the laboratory had indicated for high risk cases.

DR. FRANCIS: Dr. Birdsong.

DR. BIRDSONG: I am almost reluctant to drag this out even further, but it seems to me, if you presented me with all the data that we have been presented with, minus any reference to high risk slides, and then with reference to Dr. Bonfiglio's earlier comment, I would -- you know, my initial response would be, I would definitely want to run the high risk slides on the machine, and then perhaps manually review them, if they were AutoPap No Review, as opposed to, excluding them up front. Because, you know, if

I can believe all the data that has been presented, if anything there is an increase in sensitivity -- and so I would certainly want to apply that to the highest risk patients, if all the data is to be believed. You know, so, tell me where I am wrong there.

DR. BONFIGLIO: You are not wrong as far as I am concerned. If it was in my laboratory, that is exactly what I would do. I would not decrease what I would do with the high risk slides, anyway, but in addition, I would run them through the machine, because I think it would be additional information. And then we currently will do QC on high risk cases, and I would do that same QC, even if they were in a No Review pile. But that is not how the study was designed, so we really could not present you data in that regard.

DR. WILLIAMS: Well --

DR. FRANCIS: Dr. Williams.

DR. WILLIAMS: Well, I am just a little concerned about putting the AutoPap out to market to the private physician, and if he or she may have a high risk population, then it would be nice to know how AutoPap does in this population. And I understand you do not have the information, but it would have been nice to see out of all these people that were excluded, just exactly how well the AutoPap did, and be -- you know, just for the private

physician, so that he would know.

DR. FRANCIS: Dr. Davey?

DR. DAVEY: Are we moving on to different --

DR. FRANCIS: If it -- well, let's just make sure.

Is there anybody else who wants to address this high risk issue before we move onto other things? Hearing none, Dr. Davey.

DR. DAVEY: I wanted to make a couple of points. First, I wanted to compliment the company on a much cleaner study than the one we heard last year. I was involved closely with the Intersociety Working Group document, and I would agree in general that this study follows most of the principles.

There are a couple of things I wanted to bring up. One of them is the specificity issue, and one of them is labeling, in terms of infections and so forth.

In terms of the specificity, I think we sort of got at this earlier, but we are looking at laboratory versus clinical specificity issues, and the way I understand specificity being addressed here is more laboratory specificity; how it went to the -- what the external discrepancy panel decided, and it was not biopsy follow-up, whereas clinical specificity -- I would think what happens to the patient long term.

In the Intersociety Working Group document, it was mentioned that it would be helpful for at least a subset of cases -- and I think particularly like when you are talking about things like high grade lesion -- to know the biopsy follow-up. I mean, I know that -- I agree completely, you cannot use biopsy follow-up for all cases, because you cannot use them for the atypicals, and a lot of the low grade lesions will regress. But, I just was wondering if there was any information -- particularly for the high grade lesions and cancers -- biopsy follow-up, which would be more of a clinical specificity, instead of a laboratory specificity. So, that is one thing.

The other thing is, labeling information. I think the adequacy has been much better addressed, and I am pretty satisfied with that, but infections I do not think are a major point of Pap smear screening and I think, you know, a lot of us feel pretty strongly about that, however some people still would like to know if there are infections on the smear, and I do not think there is a lot of data on how good we are with manual screening. But, it would still be helpful to know how the instrument does, so that the user would know.

If a clinician wants to know about an infection, do they have to manually review the slide, and maybe not

rescreen it, and so do you have information on that?

DR. BIRDSONG: Can I add some comments before you answer, because that is -- almost those same comments are in my notes, too, and maybe you can answer them all at once.

Well, just in particular, in regard to infections, while I agree with everything Dr. Davey said, particularly in regard to something like herpes, which is an incurable sexually-transmitted disease for which the Pap smear is not the primary means of diagnosis.

I was just having a casual conversation with one of our gynecologists and he said about 50 percent of the herpes diagnoses that they get from Pap smears are, quote, surprises.

Well, that is still -- even though it was not being looked for, that is very useful information. And so, for the reasons she put forward, we would like to know -- I would like to know -- how the machine performs. At least in our lab -- herpes at least has a lot of reactive changes associated with it, but there is no mention of performance with regard to reactive changes, or infections.

Then comment number two, again, with regard to the Intersociety Working Group, there are a couple of -- while in general as she said, the requirements are met, there are a couple of areas that I found not quite up to that

document.

When she mentioned the biopsy follow-up, in particular with regard to the positive predictive value in terms of clinical usefulness, you know, I think that is necessary, and it would be something that we at least would want to see followed up on, even if approval is granted, as a postmarket surveillance.

Secondly, in the Intersociety document, the suggestion that ROCs be used was mentioned, and also, in some of the specific information that was sent out to the Panel -- or at least some of us -- prior to the institution -- prior to the beginning of the study, and so, we know that that issue was mentioned, and the decision was made not to present the statistics in that forum, and it would seem to me that the data presented would -- could also be favorably presented in that forum, so why was it not presented?

MS. NORTON: If I may summarize what I heard the questions to be -- and we can take them in this order, or in another, if you would prefer -- this is Mary Norton -- I heard Dr. Davey comment that she would like to see some biopsy confirmation data.

We do, in fact, have biopsy confirmation data on a subset of our HSIL+ slides that we can show you. We also have prepared data on our performance on infection, NBCC.

And I would like Dr. Wilbur to address all of those issues, then I can comment on the ROC analysis, and I would like our statistician to comment on that. And I think that addresses the issues, if I have covered them. Okay, if I could have the slide, Tim, on the biopsy confirmation data.

We have infection data that was submitted to the agency -- if I could just comment -- before we look at the biopsy data -- on infection. To your comment on herpes, we did not present that data to the agency, so we cannot discuss here, but we did collect that information.

DR. WILBUR: Basically, looking at all the cases, we attempted to find as many cases of biopsy confirmation on categories of HSIL and higher, and to date, these are the data that we have, and it is, again, incomplete.

These are the biopsy findings, illustrated across the top here. ASCUS, AGUS, LSIL, HSIL, and Cancer. And these are the numbers of cases which in follow-up had each diagnosis.

You will notice that 20 HSIL and above were in fact confirmed as HSIL and above, 6 were LSIL, there was 1 ASCUS, 1 AGUS. You will note that there were no negative biopsies that we have been able to obtain at the present time, but we only have 29 out of -- what is the total -- 70, at this point, that were available.

DR. BIRDSONG: How -- your -- I mean, you have a biopsy diagnosis in some or -- well, one at the end is ASCUS. How --

DR. WILBUR: Somebody made a diagnosis of atypical squamous epithelium, or something to that effect, and for the purposes of the study, we just put it into ASCUS, realizing that that is not a histologic diagnosis, that is a Bethesda System cytologic category, but for the purposes of --

DR. BIRDSONG: Okay.

DR. WILBUR: -- consistency, we put it in that way.

DR. BIRDSONG: I really asked that question to make sure I was reading the chart right.

DR. WILBUR: Yes. This is atypical squamous epithelium of some kind, or unspecific atypia -- an unspecified glandular atypia.

DR. DAVEY: Is there any indication that there was a difference between the current practice and the AutoPap practice arm in terms of things that were called a high grade, and how many of them had biopsy follow-up?

Do you know what I am saying? I mean, were more of the AutoPap ones confirmed, versus more of the -- you do not look at the final, but if you look at the -- each

practice arm.

DR. WILBUR: I do not think we have analyzed it in that fashion. These were compared to EDP truth, which had no bias between which side the actual diagnosis came from.

DR. FELIX: Which -- I am sorry -- which brings a quick question, who were the members of the EDP?

MS. NORTON: There are 24 of them in total, and I cannot remember them off the top. I could probably do a pretty good job, but probably not all of them.

There were 3,093 cases that required EDP adjudication, with three reads per case; so, a total of about 10,000 reads on the slides. But there were 24 pathologists. They had to be board-certified cytopathologists. So, that was the case, but there were 24 and I --

DR. FELIX: Were there a bulk of them that were read by a certain number that -- you know, your prominent readers, who are they? The people who did the most -- you know --

MS. NORTON: They were all evenly distributed among the 24 cytopathologists as best we --

DR. FELIX: And you are not going to tell me who they were.

MS. NORTON: Dr. Mark Stohler, University of

Virginia; Dr. Fauti Kareem; Dr. Michael Henry -- I am going to start losing names --

DR. FRANCIS: I think in the interest of time, we need to --

MS. NORTON: -- Dr. William Tench -- but we can provide all that.

DR. FRANCIS: Yes. If you perhaps would provide Dr. Felix --

MS. NORTON: Certainly we can do that.

DR. FRANCIS: -- with that information after the session is over.

DR. WILBUR: The next part of the question is infection, and in fact, we did do a compiled infection analysis, for which the slide should be coming up shortly here.

This is, again, infection performance, but it is combined. Actinomyces, candida, coccobacilli, herpes, and trichomonas, were all combined for the purpose. And the way they are illustrated is infection versus no infection in current practice; infection versus no infection as detected by AutoPap-assisted practice.

In this, you could see that 784 versus 940 were the discordants; with 2,141 detected by current practice; and 1985 detected by AutoPap-assisted practice. And in

fact, if you do the same statistical analysis, there is no difference with a high degree of statistical significance between those two numbers. Now, again --

DR. RENSHAW: Did you -- I am sorry. Did you do the analysis to see if they were statistically different?

DR. WILBUR: Pardon me?

MS. NORTON: I would like to have Dr. Richard Chiacchierini comment on that question.

DR. CHIACCHIERINI: No. This particular analysis was intended to determine whether or not AutoPap was statistically inferior to current practice. And the analysis determined that it was not inferior to current practice. Now, whether or not current practice was better than AutoPap was not an issue for our investigation. Okay? Does that respond to that question?

Now, I would like to address the ROC curve issue for a moment. In order to have a receiver operating characteristic curve, one has to have one or more parameters of the system vary.

We locked into a 15 percent AutoPap review only -- I am sorry -- a 25 percent AutoPap review only rate, and a 15 percent QC rate at the beginning of the study, and did not allow that to vary.

Likewise, without varying something in the current

practice arm, it would be impossible to generate a curve, as you know, because you have sensitivity versus 1 specificity, being done at those varying points of that parameter.

That was not done in this study. Those were locked in. The only other issue that could have occurred would have been a varying of the prevalence of the disease, and because we have the five centers, we thought that that would -- the pooled five centers gives you a single prevalence number.

Since there was nothing to vary, the ROC curve becomes a ROC point. And so, we really did not do that calculation.

DR. FRANCIS: I really want to bring the discussion to a close, because we are almost ten minutes over time and we really need to get to the deliberations of the specific questions before the Panel.

We will have an opportunity to address these issues a little more after the break as we get into our discussions, so we are scheduled for a 15-minute break. I want to curtail that by five minutes and ask that we reconvene here at ten minutes to 3:00.

This session stands adjourned.

[Brief recess.]

Agenda Item: Open Committee Discussion**(continued)**

DR. FRANCIS: Okay, I would like to go ahead and get started on this section of the meeting. I know that as we stopped before the break there were still I think a number of Panel members who still have questions they wanted to ask. I am going to give them the opportunity to continue asking questions of the sponsor, especially, and I will start with Dr. Renshaw.

DR. RENSHAW: Yes, thanks. I was hoping you could help me with my math. In Dr. Bonfiglio's presentation, I believe he said that the false negative rate, or false negative proportion, using ASCUS as a threshold in the current study was 21 percent, and with the AutoPap system was 14 percent -- I think those were the numbers.

When I look at the table in Volume 9 of the clinical study -- it is on page 92, it is Table 10.3 -- and then Table 10.5 -- I get different numbers; numbers of about 26 percent and 21 percent. Am I doing it wrong?

MS. NORTON: If you will give me a minute, I will get to that spot.

DR. RENSHAW: Page 92, Table 10.3 and then page 92, Table 10.5.

MS. NORTON: Those tables would give you

information about the positive predictive value related to the sensitivity, but not the complete information that you would need to do a false positive calculation.

DR. RENSHAW: Well, what about in Table 3, line 2, called, within normal limit by CP, and then you have 232 ASCUS, 9 AGUS -- so on and so forth; there are like 291. Aren't those false negatives, or am I misinterpreting it?

MS. NORTON: Those are ASCUS+ cases for the false negatives, for the current practice arm. What you might have been calculating was, just looking at ASCUS alone, you would need to combine all of ASCUS+ --

DR. FELIX(?): He is combining them.

MS. NORTON: For the false positive calculation, or --

DR. RENSHAW: Then you are telling me, the sum of the 23,556 plus 232 should equal 23,885.

MS. NORTON: No -- could you state that again -- that the sum of 23 --

DR. RENSHAW: So, you are saying the correct number of false negatives there, the total is 232, not 291, using ASCUS as your threshold.

MS. NORTON: No, it would be calcu -- it is plus the -- 45 plus 9 plus 3, so it is for ASCUS+. So, it would be the 291, that is correct --

DR. RENSHAW: Right. So 291 over your total
abnormals of 1,100 is about 26 percent, something like that.

MS. NORTON: That --

DR. RENSHAW: Or, was that -- or is that just
wrong? I mean, I may be misinterpreting this.

MS. NORTON: I think you need to look at the
bottom row for the total number of abnormalities.

DR. RENSHAW: Okay, which number would that be?

MS. NORTON: That would be the 998 ASCUS, the 51
AGUS, 278 LSIL, and the 67, 1, and 2, those would be the
total number of -- it is out of that total. Those are the
total number of abnormalities called by the truth determination
process, so it would be 1,397 is your denominator. And then
you would take the 291 over the 1,397.

DR. RENSHAW: Thank you very much.

MS. NORTON: If I may be recognized. For the
Table -- the BCC Table was requested prior to the break, and
we did not have an opportunity to present that. I just
wanted to raise that, if Dr. Davey or Dr. Birdsong wanted to
see our performance on BCC.

DR. BIRDSONG: Yes, I would.

DR. DAVEY: Yes --

DR. FRANCIS: Please show it.

DR. DAVEY: And also, if you have anything on the

different types of infections -- I mean, I do not expect statistics, but I mean, anything even just --

DR. WILBUR(?): Raw numbers.

DR. DAVEY: Raw numbers, like, is it better for like one versus the other? I mean, I know you are looking at most of the slides anyway, but that would be --

DR. WILBUR: Well, let me address this question, since we have it up, about benign cellular change performance. As you can see, again, illustrated in a 2 x 2 grid in which the top is current practice; the side -- left side -- is this AutoPap-assisted practice.

Again, you look at the discordant cases to do the test for statistical equivalence, but if you look at the ends -- on the right for AutoPap-assisted practice for number detected. And the bottom left, recurrent practice, BCC detected and the numbers are 3,431, versus 3,276. Again, there is a statistical test that was performed that is highly statistically significant.

MS. NORTON: With regard to the question of the types of infections detected, I would like to ask the FDA if we would be allowed to present our results on those categories, to address Dr. Davey's question.

DR. GUTMAN: No, you cannot introduce new information. You could provide a general description, but

you cannot produce any new data at this point. Can be presented after the Panel meeting. And if someone has a special interest, we would be happy to share it with them.

MS. NORTON: Dr. Wilbur, would you like to give just a general description of the infection performance.

DR. WILBUR: Well, I cannot present any specific data and all I can comment on is that there are no -- just based on looking at the tables that I have before me -- that there are no outstandingly different results from current practice and AutoPap, based on all of the categories that I am looking at here. That is the general answer, and if you want specifics, it is all available, after the Panel meeting.

DR. FRANCIS: Thank you. Any other general questions for the sponsors before we proceed to the questions?

DR. BIRDSONG: I had several, and these I think are more minor than the other things but I still wanted to at least mention them. Based on this, and on the previous presentations, you know, the statements have been made about the detail system integrity checks in the system, and while I am not an engineer or a physicist, it is quite easy for me to imagine very -- you know, good system integrity checks, after the image has been digitized. But I have been just

wondering, or I would like to have -- well, maybe not more detail, but you know, I have a concern about real mundane things possibly messing up the procedure.

For instance, you know, the old proverbial, you know, dirt on the objectives, or, you know, if a focusing motor fails -- would things like that also cause a trade to get rejected?

MS. NORTON: I can comment on that. Yes, that is in fact the case. The instrument is highly sensitive to dust, and to any other mechanical or features that might be present on the slide that would prohibit it from being processed. It would be rejected in the same manner as any other situation.

DR. BIRDSONG: Okay, another question -- you can stop me if I have too many of them, but with regard to the laboratory process qualification procedure. Based on reading of all the information, it seems like to some extent the machine is calibrated to the baseline performance of the lab.

At the same time, obviously, there are going to be limits there. But I was concerned if -- is it possible, let's say you are setting up the machine in a lab that is performing marginally to start with, is it possible for the slide score thresholds to be set inappropriately, if a lab

has a very high false negative rate to start with? Is that question clear?

MS. NORTON: No, that would not be the case. They are independent of whatever the false negative screening rates of the laboratory are. They are set based on a sample of the slides that the laboratory has recently assessed. So it is independent of both --

DR. BIRDSONG: But you are not looking at the lab's diagnoses --

MS. NORTON: That is correct. It is only measuring the properties of the slide, as it is submitted to the AutoPap for processing.

DR. BIRDSONG: Okay --

DR. FRANCIS: Any more questions, Dr. Birdsong?

DR. BIRDSONG: No.

DR. FRANCIS: Dr. Davey.

DR. DAVEY: I have more, as usual. I wanted to just cover a little bit more about what Dr. Birdsong brought up about maintenance sorts of issues.

I understand a lot of this is sort of electronic quality control, and if there is something that changes, you are basically recalibrating, like you would a hematology instrument, for example.

Now, what about, is there any sort of routine

quality control that is possible? I mean, you know, we are used to sticking through known normals and known abnormal, for example, in a hematology instrument. Is that something that could be easily done just to satisfy the customer that -- just would feel better about that, and is there anything that you would routinely recommend, initially? So, that is sort of one question.

Then, another issue, too, is when we had the Panel last year, we had variation in performance in different laboratories; I think some of the university laboratories had lower false negative rates to begin with, and I think that the laboratories that were used for this prospective trial are sort of different.

Would you see any differences in different types of laboratories in which you introduced this now, or -- for example, a university laboratory that has a lot of high risk patients, would that make a difference in how you would set up the instrument?

MS. NORTON: We will take your second question first, and I would like Dr. Bonfiglio to respond to that question, and then I will respond to your first question on the quality assurance processes.

DR. BONFIGLIO: I do not think a university laboratory would have any different set-up of the instrument

than any other laboratory. In regard to whether performance would be any different, I think -- obviously, the performance, or how many cases -- false negative cases -- the instrument is going to detect depends on what the laboratory's underlying false negative rate is to begin with.

In a laboratory that has a higher false negative rate, obviously, the instrument will pick up a higher proportion of cases, and therefore, improve that laboratory's performance more, say, than a laboratory with a lower false negative rate. However, I cannot envision a situation where the instrument would not improve the performance of the laboratory, unless the laboratory had a false negative rate that was lower than the false negative rate in the AutoPap alone screen category, which is -- at least in this clinical study -- less than 3 percent.

As we know, I think most of us in the field have pretty much agreed that there is -- with the current technology, there is an irreducible false negative of somewhere between 5 and 10 percent, so that no laboratory that I am aware of is functioning at a rate that is below that 3 percent.

DR. FRANCIS: Dr. Birdsong --

MS. NORTON: Your second question, Dr. Davey -- I

am sorry --

DR. FRANCIS: Go ahead.

MS. NORTON: -- was with regard to what recommendations we might have for laboratories that wanted to, say, process additional samples? Certainly, we would consider that that would be an option of the laboratory to do that, and in fact, I think it is true that current customers do that as a matter of course, anyway; that they select a certain number of slides that they routinely run on the AutoPap instrument, so I certainly would say we would not be -- we would be open to that, you know, as a concept, and I believe it is going on today in the laboratories that currently use the AutoPap instrumentation.

DR. FRANCIS: Dr. Birdsong?

DR. BIRDSONG: Another question. I think -- you know, maybe for Wilbur, but -- or anyone who has had experience. Has there been any characterization of the acknowledged false negatives that -- few though they may be -- the false negatives that fall into the AutoPap No Review category, just -- you know, running a lab, you like to know what the capabilities and potential shortcomings of any piece of machinery you use are, and you know, even if your overall performance improves using that, it would be nice to know, are there any identifiable characteristics of false

negatives, or would you describe them as just random misses?

DR. WILBUR(?): Well, we have not studied that aspect, particularly, but it is certainly something that would interest me, as it does you, and I am sure that at some point, we will study that, but at the present time, we have no data on that. We have not had a chance to specifically analyze those cases.

DR. FRANCIS: Dr. Felix.

DR. FELIX: In the pamphlet that was presented to us by the FDA statistical team, we find that at the conclusion of this study, an average of 20.25 percent of the slides analyzed were for No Review, whereas 17.86 percent were for QC review.

The number that -- the goal that you are proposing is of a 25 percent range, and I find a significant -- or at least at a casual glance, a significant difference. Can you reconcile these numbers?

MS. NORTON: I would like to comment that, from our claims, that up to 25 percent of the slides could be classified for AutoPap review only. And that does vary on a day-to-day basis, based on the prevalence of disease on any given day or time period that is being measured.

It will never exceed 25 percent, but if you remember, site-by-site, we actually had one laboratory whose

AutoPap review only rate was 24. -- I believe -- 56 percent -- I could be off by a few tenths there. And that the No Review population is for the whole population, including the abnormal slides. You have to keep that in mind, that that does vary.

DR. RENSHAW: Does it include the slides that are not able to be run on the machine, as well?

MS. NORTON: No, it does not.

DR. RENSHAW: So, really it is -- since about 20 percent of the slides cannot be read on the machine, it is really less than that, even.

MS. NORTON: It goes from approximately 25 percent to 20 percent.

DR. RENSHAW: We are in this average from 20 percent to 16 percent. Something like that.

MS. NORTON: If you take away the slides designated as process review, is that what your comments are --

DR. RENSHAW: Or if you include all the slides in the lab, the average in this study was much closer to 16 percent than 25 percent.

PARTICIPANT: Because of the exclusion, it is about the excluded --

DR. RENSHAW: Since about 20 percent did not even

--

MS. NORTON: Right. That is right. If you were to look at the slides that were excluded, right, the ones that were not run on the instrument.

DR. FRANCIS: Dr. Williams.

DR. WILLIAMS: I had a question for any one of the Panel. The slide that you showed us for indications for use, being that you talked about 20 percent or so of the slides that were being excluded for one reason or another, would it be more accurate to put down that this device should be used for low risk and quality control rescreening, instead of primary rescreening? Would that be more accurate, in this case, since we are excluding a lot of the high risk patients?

MS. NORTON: I would like Dr. Wilbur to comment on that.

DR. WILBUR: Well, the only thing I would comment on is that would be at least in the current jargon, not really in synch with the way things are categorized. Really, we specifically exclude high risk populations, but we do not specify what constitutes a low risk population.

Now, certainly, it is one minus the high risk population intuitively, but that is not certainly a common jargon that I would at least think would fit.

DR. WILLIAMS: Well, what would you think would fit, in this case, since we are excluding, quote, high risk patients, what do you think would be more accurate? Because if you -- can you flip that slide up, when you put down the indications for use? Is that your slide, or --

PARTICIPANT: -- FDA's --

DR. WILBUR: I think that was FDA's.

DR. WILLIAMS: Can you flip that slide up for a minute?

DR. WILBUR: -- see that that is relevant, I do not understand what -- this is just -- rename it.

DR. BONFIGLIO: There was something I was thinking of, but then I just lost my train of thought, and --

DR. WILLIAMS: See, I have a question about, since we are prescreening the high risk people out of this system, don't you think we should mention something about the indications for usage of this machine?

DR. BONFIGLIO: That is labeling, isn't it?

DR. WILBUR: That is a labeling question, and I am sure in the --

DR. FRANCIS: That is going to be covered in question two.

DR. DAVEY: I have one more -- I am sorry, anybody else? One more quick question. I think that a new

algorithm -- am I correct in understanding that there aren't abnormal cases that -- you know, if you look at the old, historical sensitivity studies, we should just be gaining sensitivity. Are we losing any cases? I mean, will there be any cases like, some of the more rare conditions that would be not missed, that would be not picked up by the new algorithm?

MS. NORTON: No. In fact, the existing gen IB algorithm as in the current approved product remains intact in the new instrument. As Dr. Nelson showed, what you do is you add an additional piece of information at the AutoPap review, only, and the quality control rescreening, to enhance performance. So, it cannot degrade performance, only improve it.

It is important to note that, at least 15 percent of the slides are selected for quality control review, so if that were to go to 17 percent, that is a good thing; more slides are being selected for quality control review; that is not a negative situation. But in the case of the AutoPap review only population, it would not be allowed to exceed or go higher than 25 percent AutoPap review only rate. It would only be allowed to decrease; meaning, more slides designated for AutoPap review.

DR. FRANCIS: Dr. Birdsong has a comment.

DR. BIRDSONG: So, under no circumstances would fewer than 15 percent fall into the QC review?

MS. NORTON: That is correct.

DR. BIRDSONG: Okay, now another question --

DR. FELIX: Can I just chase one second, because it was a question that I had been meaning to ask --

DR. FRANCIS: Dr. Felix?

DR. FELIX: Thank you. In this study, the AutoPap decided that 20 percent was the most that it could classify as within normal limits. And yet, looking at the numbers of the population that you studied, there was in fact a very low number of abnormalities. I think it was 1 percent low grade, .3 percent high grade. How would you envision doing in a population with a more high rate of abnormalities? I mean, is it going to go down -- what is the percentage of slides that AutoPap is going to say, these are negative?

MS. NORTON: The instrument -- the AutoPap -- could select up to 25 percent. The 20 percent represented the average during this duration of the study. So, presumably, that could vary up towards -- and as in the case of the one lab -- it varied up to 28.8 percent, so that will vary on a day-to-day basis. So, that was the first comment.

To the second comment on the prevalence in the study; in fact, there were 1,397 abnormal cases, which is

approximately 5.6 percent of the population of slides in the study, which I would let Dr. Wilbur or Dr. Bonfiglio comment on further --

DR. FELIX: That is including ASCUS, though.

MS. NORTON: That is correct. For all abnormal.

DR. FELIX: That is pretty low in today's standard, am I in agreement here?

DR. BONFIGLIO: Yes, but remember, you excluded the high risk patients.

DR. FELIX: Oh, okay.

DR. FRANCIS: I will take two more questions before we really do need to move to discussion of the specific questions before us. Ms. Rosenthal?

MS. ROSENTHAL: Something that occurs to me is that, if we did have the high risk patients in this study, probably it would look that much better, because, they would show up.

I wonder if, in the clinical -- in the real world -- we really want to keep high risk patients from being run through this system, or do we want them actually to be included in this system, because they may have a better chance of having, you know, an event detected?

DR. BONFIGLIO: Yes. I would like to comment on that. I -- and I think I mentioned it when I said how I

personally would use the machine --

MS. ROSENTHAL: Right, you did.

DR. BONFIGLIO: And for high risk patients, I certainly -- I would run them through, and then I would do what my laboratory does with high risk patients, anyway. So, I think you are right. But I do not think we -- I mean, I am sure the company cannot claim that, because we could not design a study so we could include high risk patients in the study.

I think that most laboratory directors, once they understand this, will use it in that way.

MS. ROSENTHAL: Might we have some continuing research, to just -- or some continuing trials, to see what it does look like when we use high risk patients, also?

MS. NORTON: I have just a follow-on comment, and then perhaps I could answer that question. I would like to mention and remind of the statement I made earlier that, in fact, in previous studies that were submitted, we looked at our performance on both high risk and non-high risk cases on the large number of disease categories. In fact, there was not any difference in terms of the AutoPap's assessment of those high risk categories. So, I think, given that that baseline of performance is resident within this version of the AutoPap, we would not expect to see that; however, at

this point in time, the company is not seeking any claims for the high risk running of slides on the AutoPap. And it remained as a limitation on the quality control instrument, as well.

DR. FRANCIS: Does anybody have one more question?

Dr. Allen?

DR. ALLEN: I do not run a lab, I am a clinician, but in the data presented by the FDA, about 20 percent of your slides could not be processed by the AutoPap method. How does that compare with the standard methodology for clinical review; what percentage of slides are unreviewable clinically? Is this higher, or lower, or about the same?

Then, does this automatically -- the slides that cannot be processed -- I guess half of them were high risk patients, and the other half for technical reasons. Does that then go to your QC arm, that require manual review, which will bring up the percent of QC from what was 15 percent? Would you add that 10 percent, and then make it -- that half of these are technically unprocessable, then they would go to the QC arm. It actually adds the percentage that should be in the QC arm, that realistically will be -- and then, again, to repeat the first half of the question.

MS. NORTON: Dr. Wilbur, would you like to comment on this issue?

DR. WILBUR: I think if I understand the question properly, what you mean is, if you exclude these slides, what in fact happens to them?

DR. ALLEN: Right.

DR. WILBUR: Well, I mean, it would be exactly what would happen to them, would they not be subjected to the AutoPap-assisted practice protocol. For instance, if you excluded a slide because, say, it was broken. I mean, you would do your best to fix the slide, and then of course it would have to be manually screened, because in fact, the AutoPap could not process it.

Under those circumstances, if in fact there was something identified that was abnormal, it would go to the pathologist. If it fell into some risk category, either defined by your laboratory -- meaning it was satisfactory but limited by your current laboratory practices, that those are automatically QC'd, then it would be QC'd.

If not, and you did not meet any criteria for quality control, then it could essentially be signed out as negative. And that goes for any such exclusion category; the high risk would just follow your own protocol through. And in fact, the point is, it would not add quality control, because now you are looking at discrete blocks of slides; you are looking at, for instance, 15 percent of x-number of

slides, and now, 15 percent or 10 percent, or whatever your QC is on the exclusion of another block of slides. So, your overall percentage does not increase, it remains the same.

DR. ALLEN: And how does this percentage of slides that cannot be reviewed by this technology compare to the manual methods that we can point out?

MS. NORTON: I think maybe a point of clarification might help answer that question. The reason why slides are rejected by the AutoPap is not because the AutoPap is not processing the slides, it has to do with the fact that there is something wrong with the sample slide as it comes into the instrument; meaning, the staining has varied. We take very close measurements of how staining might vary. If coverslips are tipping. If there are chipped edges. If they are of unusual shapes or sizes. If there are bubbles under the coverslip.

All of those things are measurements that the AutoPap is making, so the reason why a slide is rejected is because the laboratory's process may have drifted. So what we often find is that, when we are able to give that information to the laboratories by rejecting the slides, they are able to improve their processes or tighten up their limits, so it does not really -- the comparison between the manual practice and the AutoPap performance, I do not know

whether that is a question that I can answer that might clarify that issue for you. I do not know if it does or not.

DR. ALLEN: Okay.

DR. FRANCIS: Dr. Birdsong?

DR. BIRDSONG: This is a statistical question, and I returning to something we discussed earlier on the ranking system. I was looking at the numbers and actually computing some of the ratios for AutoPap gain versus CP gain. And it appears that, essentially all the trend, if you will, is established essentially by group 1.

If you look below group 1 for ASCUS+, or LSILs or LSIL+, as Dr. Davidson was mentioning, it appears relatively level, or you know -- but when you go up to group 1, then there is a big increase in the AutoPap gain.

Now, I am not a statistician, but is the way I am reading this right? Is, basically, all the gain -- I mean, the statistically significant trend that was noted in the data, is that essentially all due to the fact that the AutoPap gain is so much higher in group 1 than it is in the others? This is Table 10.20 I am looking at.

MS. NORTON: I would like to have Dr. Richard Chiacchierini comment on that.

DR. CHIACCHIERINI: Your observation is quite

correct in the sense that, the highest 20 percent of slides ranked is where you would like to find the most pathology, and in fact, the statistical test of the ranking system, first the discovery, was in fact statistically significant? But I need to point out, this was not a study that was designed to evaluate the behavior of the cytotechnologist, so we cannot say anything further than that. All we can say is that there was a strong correlation between the ranking - - the slide ranking -- and the pick-up of the pathology.

DR. BIRDSONG: Right. Now, if for instance, if we did the same statistical test, but removed group 1 -- you know, eyeballing, it looks like there probably wouldn't be a significant trend, is that correct?

DR. CHIACCHIERINI: It may not be. We did not do that test.

**Agenda Item: Panel Vote and Recommendations to
FDA**

DR. FRANCIS: Thank you very much. I would like to move on now to discuss the specific questions before the Panel, and what I propose to do is read aloud the question for the benefit of the Panel and for the public, and then I propose to go around the Panel individually, polling them for their responses to each question.

After each member of the Panel has given their

thoughts and comments on each of the questions, I will invite the sponsor to comment on any of the comments made by the individual Panel members.

The first question for Panel deliberation you will see on your screen momentarily. Do the data presented in this PMA support the manufacturer's intended use of the AutoPap System? And I will start with Dr. Nestok.

DR. NESTOK: I would say that, from the data presented today, as compared to the initial meeting we had on this PMA, that, yes, that the data do support the intended use of this system.

DR. FRANCIS: Thank you very much. Dr. Davey.

DR. DAVEY: Yes.

DR. FRANCIS: I like those comments. Dr. Williams.

DR. WILLIAMS: Yes. I have some concerns about the high risk patients, but, yes.

DR. FRANCIS: Dr. Renshaw.

DR. RENSHAW: I have two concerns, one that there were not enough HSIL patients to be statistically significant. And I do not think photocopying is the best way to preserve the things; it should have been a cross-over study. But with those two limitations, I still have to say yes.

DR. FRANCIS: Thank you. Dr. Felix. I missed one. Dr. Allen, I am sorry.

DR. ALLEN: That is quite all right. Yes.

DR. FRANCIS: Dr. Felix.

DR. FELIX: Yes.

DR. FRANCIS: Ms. Rosenthal.

MS. ROSENTHAL: Yes.

DR. FRANCIS: Dr. Floyd.

DR. FLOYD: Yes.

DR. FRANCIS: Dr. Singh?

DR. SINGH: Yes.

DR. FRANCIS: Dr. Birdsong.

DR. BIRDSONG: I would have to say yes, with the caveat that, actually, even though the study was not -- the study was designed with the high risk patients excluded, nevertheless, the data in this presentation and from their previous presentations, the data suggests that they should be included, I would think. And I also would agree with Dr. Renshaw. I would have liked to have seen more HSILs in the prospective study, but nevertheless, I still think the data support the intended use.

DR. FRANCIS: Thank you. Dr. Davidson?

DR. DAVIDSON: I do not think I am included in this part of this --

DR. FRANCIS: You can just --

DR. GUTMAN: We are not at a vote yet for approval or disapproval, so you can certainly comment on it.

DR. DAVIDSON: Then I would just say, yes.

DR. FRANCIS: Yes. Thank you. Is there anything the sponsor would like to comment on? I think it was --

MS. NORTON: I do not think we have any comments.

DR. FRANCIS: I think it was fairly unanimous.

Now, the second question before the committee. Are the claims being suggested by the manufacturer, will their device's performance compare to manual screening, supported by the data available? If so, how should these claims be presented in the labeling and what, if any, limitations should be applied to this data presentation? Dr. Nestok.

DR. NESTOK: In short, I would say, yes, but it does bring up one question that I did have, and it probably should be mentioned in the labeling and that is about benign endometrial cells. Could someone comment about the device and the ability to pick up benign endometrial cells, and how that will affect practice with this device?

DR. FRANCIS: Perhaps if you would -- note that question, Sponsor, I will come back to you at the end, and you can address that issue. Dr. Davey.

DR. DAVEY: Again, I would say, yes, I do think we

need to talk about some labeling and maybe as well as some maybe postmarket surveillance on some of the things -- I mean, I think the infections we probably need to bring that up in the labeling, depending on what data there is, because I think that -- you know, again, I do not think that the instrument needs to prove anything, but it just needs to be obvious to the user how the instrument does.

The benign cellular changes, I am not as concerned about because I think that is a waste -- a bit waste basket category, and I am sure that it is clinically -- what the clinical significance is, that if you looked at that it is, you know, it is -- the inner-observer agreement for that is even worse than ASCUS. So, I think the infections we need to know, and I think maybe it would be of interest to know how the instrument does in more high risk populations afterwards, and with high grade lesions. So, those would be the things I think maybe increase some of the information available to the user.

DR. FRANCIS: Thank you. Dr. Williams.

DR. WILLIAMS: I have to agree with Dr. Davey. I am, you know, concerned about the clinician out there in the real world that has to treat patients and a significant amount of their patients may have infections -- may have had an abnormal Pap in the past. You know, it is just a lot of

exclusion, but I would still have to say, yes, with the same reservations.

DR. FRANCIS: Dr. Renshaw.

DR. RENSHAW: I would say, yes, with the caveat, in the labeling, it repeatedly says that, up to 25% of the slides can be classified as No Review. That excludes that 20 percent that cannot be processed. It should be based on all the slides in the lab, so it should be about 20 percent, and I would also request that they actually report an average as well as an up-to, to give the people who use the machine a better feel for what kind of benefit they actually may get.

DR. FRANCIS: Thank you. Dr. Allen.

DR. ALLEN: I agree with everything that has been said to date, that the claims that have been suggested have been supported by the data. I would like that the limitations be spelled out. I do not know -- as a clinician -- rare events, like tubal carcinoma are often picked up on the Pap smear, so you did not look at that, if that just -- you know, it is a rare event that the device has not been tested in high risk populations, and I personally practice in a university setting in a particularly high risk population, and would love to have a test that improved sensitivity. Because those are the patients who need the

lower false negatives. They really need the true positives.

That limitation -- you know -- that the labeling be clear as to what was not identified, and what populations were not studied.

DR. FRANCIS: Okay, Dr. Felix.

DR. FELIX: I also think that the claims are supported by the data, and I think everything that I have had to say has been mentioned already.

DR. FRANCIS: Ms. Rosenthal?

MS. ROSENTHAL: I think the claims are supported by the data, and I, too, would like to see some post-approval data about high risk populations.

DR. FRANCIS: Dr. Floyd?

DR. FLOYD: I agree. The claims are supported by the claims are supported by the data presented to us. I think that some of the issues raised about labeling are quite valid, and I think all manufacturers are interested in making appropriate claims for their devices, in conjunction with FDA guidance on those issues.

I also think some of the issues that we discussed here regarding postmarket surveillance may be taken slightly out of context, because postmarket surveillance is a little bit different from what is going to happen with this approved device, because a lot of people are going to use

this device on high risk populations to publish papers. So, this data is going to become available in a standard way in which medical science has advanced over the years, and I think the data presented support the claims that have been made for this device, and that is pretty much sufficient, it would seem to me.

DR. FRANCIS: Thank you. Dr. Singh.

DR. SINGH: My answer is, yes, as well, but again, I am not a lab person, and I would still say in the labeling we have to be careful, as Dr. Allen mentioned, all the three points. To say, yes, the high risk population was not studied, and also the infection issues, a lot of the times we as clinicians get that from our lab, which is very helpful to us in day to day life. And I think the limitations should be spelled out very clearly, because people just take it and run with it, and an everyday clinician does not have the time to figure out whether or not you have gone from a manual system to an automated, so I think it is very important -- your labeling is good. Thank you.

DR. FRANCIS: Dr. Birdsong.

DR. BIRDSONG: I would say, yes, I agree with the things that have been said, I nevertheless want to reiterate that the labeling should indicate that, you know, data was

not submitted to the -- well, I guess it has been submitted, but not presented to the Panel, at least, regarding infections, or perhaps that data, you know, could be submitted, you know, between now and final approval, but there should be some explicit labeling regarding the performance of the machine with regard to infections; in particular, herpes, and as I have said a couple of times before, I think that even though the study did not look at high risk patients, the data suggests that they may benefit, and perhaps some mention of that could be mentioned in the labeling, as well as the postmarket surveillance.

I think that, although, as Dr. Floyd pointed out, people will do studies and this data will become available in the standard fashion, and I think -- and you know, it is probably to everyone's benefit, you know, including the company's, to have that data explicitly available, particularly if at some point in the future, the question of a false negative that fell into the No Review category came up, it would be nice to have at least one set of data that very explicitly said, the machine is at least as humans at detecting those cases.

As far as the immediate labeling, I think we would have to mention that high risk patients were not studied, although the data suggests that the machine is effective.

DR. FRANCIS: Thank you. Dr. Davidson.

DR. DAVIDSON: I would just raise a question, rather than taking a position one way or the other. I would -- and it has to do with the discrepancies -- and whether or not there are any qualitative differences between those discrepancies, or the differences are such that that could just be ignored as just a numerical figure.

I do not know when the discrepancy occurs, whether or not that is more important group of cases in the AP group, or whether it may or may not be a more important group of cases in the routine laboratory.

DR. GUTMAN: Can I --

DR. FRANCIS: Dr. Gutman.

DR. GUTMAN: Yes, I have a question. It was actually the kick-off on the Panel today, and it only obliquely refers to labeling, but I cannot let you out of the room without at least making sure I have brought comment on this issue.

It is the issue that was raised in the public comment by Trylon, about whether in fact as a label for this type of product -- not this company in particular, but this type of product -- ought to lead into a claim for primary screening device, or whether it ought to have some other kind of claim; a Pap smear reader, a Pap device, a slide

screening device.

I got a fairly strong sense from at least two members of the Panel of their view of -- or, I thought I got a sense of their view of the comment from Trylon, but I would be curious if anyone else had a passion for the semantics of how to properly label the heart and soul of what a Pap smear would be.

DR. FRANCIS: Dr. Birdsong.

DR. BIRDSONG: If the machine is going to designate some cases as negative, it is a primary screening device, I mean, you know, period. So --

DR. FRANCIS: Anything else --

DR. FELIX: Agree.

DR. FRANCIS: -- Dr. Felix -- Dr. Davey --

DR. DAVEY: I think it is nice to put in there the cervical-vaginal cytology, or Pap smear, because -- I mean, I think the point is is that there are other ways of screening populations, and this is -- but if you are looking at Pap smears or cervical cytology, then it is a primary screening device. Agree.

DR. FRANCIS: Dr. Williams?

DR. WILLIAMS: I disagree. I just -- and I am a clinician, and I feel that if you put down there that it is a primary screener, then that is what the clinician is going

to see it as, a primary screen. So, I would say that, no, maybe this is not a primary screener, because you are prescreening some patients before you even use the machine. So, in that case, I would say that, no, it is not a primary screener. Just my opinion.

DR. FRANCIS: Other comments from the Panel on that issue? Dr. Felix -- Dr. Birdsong?

DR. BIRDSONG: Well, my comment is -- I think we are indeed playing semantics here, because -- I use the term primary, when I -- I obviously have pretty definite feelings about it.

When I use the term, primary screener, I am referring to, within the lab, you know, primary screening. And you, I think, are not referring to in the lab, when the slide is in the lab, if I can run it through a machine and it is going to be put in the file, based on something the machine says, that is the only screening it is getting in the lab; it is a primary screening.

Now, it may be that, say, the high risk, you know, gets some additional screening, too, after that -- primary, meaning first, still, you know, it is a primary screener.

You know, there are other methodologies, as Dr. Davey referred to, so, say, primary Pap smear screening device, but I don't know -- I would like to hear your

response to my response.

DR. FRANCIS: Dr. Williams.

DR. WILLIAMS: Dr. Birdsong, I agree in that instance. And I just -- I still feel pretty strongly that if you are not screening everybody, and you are selecting some population that you are seeing in other populations, you are not using the machine. Then, in my view, that is not primary screening --

DR. BIRDSONG: Then it is a --

DR. WILLIAMS: -- but, I am not a lab person. I am a, you know, a clinician, also, so --

DR. BIRDSONG: This is a -- all right, well --

DR. WILLIAMS: -- I think it is a semantic type thing. And I think it has to be solved, how we are going to do this, and since there is some type of confusion here, I think before we leave today, we have to come to some type of language that everybody is clear on.

DR. BIRDSONG: All right. Well, how about if it would say, a primary screening technology that is not applicable to every slide, or every case -- and then -- because of the data, the nature of the study that was done to validate the performance of the instrument. It is -- you know, I think it is still a -- it is a primary screening device --

DR. WILLIAMS: Well, I -- I think --

DR. GUTMAN: You do not have to come to consensus. We are just asking for input at this point, and you are more than welcome to provide input following the Panel. This is an interesting semantic issue, but I do not wish to bog down the Panel, either. I apologize.

DR. WILLIAMS: But, I will say -- I will say this, though. I think you can say in the language somewhere, that it can be used as a primary screener.

DR. FRANCIS: Thank you. Do the sponsors have any comments on anything the Panel members have said in response to this question?

MS. NORTON: I have three comments. The first comment is with regard to the discussion of infection, and BCC performance. We have provided those data to the agency. We would be happy to include that data as part of our labeling, if it was determined by the Panel that that was relevant performance information.

Secondly, that with regard to the comment about the 20 percent versus the 25 percent from Dr. Renshaw. The 20 percent numbers referred to the cases that were excluded from the study, comprised of the high risk and those cases that were not processed on the AutoPap. That is a different percentage.

The 25 percent refers to, once the slides are processed on the AutoPap, that up to 25 percent of those run on the instrument would be selected for AutoPap review only, so I wanted to make that clarification.

The third comment was with regard to high risk and that would be the subject of discussing labeling, and as I had mentioned before, we did have data in previous submissions on the performance in high risk cases, as well.

Dr. Wilbur, do you have anything to add?

DR. WILBUR: Well, I think there was one other comment that I would just like to make, and that is regarding the number of HSILs in the study, and I would like to comment that this is a prospective, intended use study, and the number of HSILs in that kind of a study is never anticipated to be huge. And that is the reason we have the supporting documentation on over 600 cases of HSIL, which I think is very similar to the study, and therefore, I think one can extrapolate the performance to a larger group.

DR. FRANCIS: Thank you. Dr. Nestok, I think you had a particular question in your comments. Was that answered?

DR. NESTOK: Yes. Could you address the benign endometrial cells situation?

DR. WILBUR: The subject of benign endometrial

cells really boils down to a couple of things. First of all, the machine was tested on grade 1 endometrial adenocarcinomas, and there is substantial data in the submission already about the sensitivity to that particular area.

I would like to refer to one of our site investigators to comment on this, because they have specific experience. Dr. Tench?

DR. TENCH: I am Dr. William Tench, Director of Anatomic Pathology at Palomar Health System, and I am an Associate Professor of Clinical Pathology at the University of California, San Diego.

We have had the device in our laboratory for approximately two years, and for a period of time, we were interested in identifying the characteristics of cases that were sort as QC review, in which we were unable to identify a significant squamous or glandular lesion.

We looked at approximately 1,000 of those cases, and found that in approximately 5 percent, the only abnormality that we could identify was the presence of normal-appearing endometrial cells.

We know that the device has certainly some sensitivity. The problem that we have, obviously, is that we do not have a denominator for that number; we are looking

at those cases that have already been sorted.

The other issue, however, that we have with that is, I think a change in the clinical attitude about the significance of endometrial cells in the postmenopausal patient. Even within my own practice, we have a tremendous range and variety of patients being treated hormonally, and that obviously has a major impact on the significance of that data that were generated in the years before postmenopausal supplementation were available. So, I think that we even have to call into question the significance of that finding in the first place, although in my laboratory, we do still report it.

DR. DAVEY: Yes, could I --

DR. FRANCIS: Dr. Davey?

DR. DAVEY: Yes. One comment on that. I think we would all have to agree -- I want to, you know -- in relation to some of the public comments this morning, we cannot prevent all cervical cancers, and particularly, adenocarcinomas. I do not know if there is any proof that Pap smears have helped decrease adenocarcinomas, and certainly, endometrial cancer, so I guess it is not -- although we would like to pick a lot of them up, it is not -- I agree with some of what has been said. It is not as huge a concern for me.

When we are talking about labeling, this probably is not really part of labelling, but I wanted to make this comment, anyway; it is more advertising that I do want to, again, make sure -- I have made this statement before -- that the company is not permitted to claim a new standard of care in its advertising. That is for the profession to set over time, and I just want to go on record as saying that.

DR. FRANCIS: Thank you, Dr. Davey. To move ahead to the third and final question for Panel deliberation. In the intended use study, the AutoPap system provided a ranked report to all the screening cytotechnologists for all review slides. Will having knowledge of the ranking affect the cytotechnologist's vigilance? Dr. Nestok.

DR. NESTOK: I think it most certainly will have an effect on the cytotechnologist's vigilance, having this ranking, although I think that, I guess you could look at it two different ways, but I think that it actually would have a positive effect in knowing the ranking.

DR. FRANCIS: Thank you. Dr. Davey.

DR. DAVEY: Yes. Working with technologists for a long time, I think it will affect their vigilance, and I am not sure -- you know, I think it would generally be positive, unless you are one of those unfortunate patients that has an abnormality that falls into a lower higher

ranking, or the ones that are not as likely to be looked at. So, I think this would be another area that would be very interesting for studies later on, because, you know, in a way, I was thinking, initially, that you would want to initially not have the technologists know, until you got to the QC part, because you would want to have everybody that was not kicked out, have an equal chance of being screened the same way, to begin with.

I am not sure if that is right or not. I mean, I can see advantages for each, because it looks like, clearly, part of the reason that the sensitivity is improved is because the cytotechnologists know about that. So, again, with -- you know, I think that is an area for additional study, but I am not sure that it really affects, you know, approval status or anything.

DR. FRANCIS: Thank you. Dr. Williams?

DR. WILLIAMS: I also agree with Dr. Davey. I think it will have some effect on the vigilance of the cytotechnologist, but I am not sure if that is important as far as approval is concerned.

DR. FRANCIS: Dr. Renshaw?

DR. RENSHAW: Of course the answer yes, everything affects vigilance, even the fact that it is raining outside. Whether it will have a positive effect is less certain.

I think the data that it performed better, suggests that it does, but it is probably the weakest or most poorly-studied aspect of the study.

DR. FRANCIS: Thank you. Dr. Allen?

DR. ALLEN: I in fact felt that I had no data with which to answer this question. If you say -- if you look at the ranking, and if you look at the fifth quintile compared to the first quintile, there is a lower pick-up in the fifth quintile. Does that mean that there is less vigilance in the fifth quintile? I do not know.

I think that it is a different kind of a study than just collecting numbers. It is almost a behavioral study that -- or a psychological study -- that would have to be done, and the data was not presented to answer this question.

DR. FRANCIS: Thank you. Dr. Felix?

DR. FELIX: The answer is of course, yes. And my interpretation of the data is that it is the only explanation that I can think of why the device aided the detection of lesions over the regularly-screened population.

If there is another explanation for the increase in that population, I am not quite sure what it would be.

DR. FRANCIS: Ms. Rosenthal?

MS. ROSENTHAL: My answer is yes, but I am

concerned that there would be more vigilance for the higher quintiles, and less vigilance for the lower quintiles.

DR. FRANCIS: Dr. Floyd?

DR. FLOYD: I would also answer, yes, and I would say that, based on my experience, anything that helps prod the cytotechnologist to give an adequate review to every slide, is to be applauded.

I would also say that this offers the opportunity to do some very interesting studies simply by flip-flopping the rankings, and see what happens to the ratings. And just as a comment, practical experience. I know that the cytotechnologists that I have supervised over the years -- because I actually surreptitiously recorded the time they spent on slides, and the amount of minutes per slide decreased over the period of the day.

The other interesting observation is that, the amount of time spent per patient depended upon whether or not they recognized that patient's name as one they had seen before. So, they spent more time on patients who were repeaters from a particular practice service, than they did from those that were an unusual name.

I think this is a very interesting question. It gives us a tool to do some very interesting studies, but I do think that it will certainly affect the surveillance of

that high ranked category.

DR. FRANCIS: Thank you. Dr. Singh?

DR. SINGH: My answer is yes.

DR. FRANCIS: Dr. Birdsong?

DR. BIRDSONG: My answer is dichotomous. As a lab director, you know, looking at slides myself, working with cytotechnologists, my answer is, yes. Looking at the data, I do not think the -- I think that yes is an answer that is consistent with the data presented, but the data presented do not convincingly prove that that is the reason.

The function of the machine is to -- or a function of the machine is to give slides scores that are proportional to the likelihood of there being an abnormality. So, another explanation of the Table 10.20 is that the AP gain versus CP gain is simply because there are more abnormal in group 1 -- I mean, that is what the machine is designed to do, and it is possible that vigilance is equal across those five categories, but there are more abnormal in group 1 as compared to group 5, therefore, you --

There is nothing analogous to the ranking given on the CP side, so that may explain it, but you know, as has been mentioned previously, this gives us a tool to do very interesting studies and that hypothesis can be tested in a

study that is designed specifically to look at vigilance. It is certainly consistent with the observed data, but the data per se do not address the question directly, and do not answer it. So, subjectively, yes; objectively, you know, I am not sure from this data.

DR. FRANCIS: Thank you. Dr. Davidson.

DR. DAVIDSON: I do not have any questions that the ranking, if it is characterized as proposed, will certainly affect the behavior of the cytotechnologist. But I am not convinced that the relative risk in any category is lessened, but I am not convinced about those numbers.

The risk may be -- though the numbers may be smaller, the relative risk, or the percentage of cases out of those smaller numbers that are abnormal, may be just -- the risk may be the same. I am not convinced of that, yet. But I am convinced, if you tell the cytologist that category 1 is of greater concern than category 5, then, yes, I have no question in my mind that that will influence the cytologist, but I would hate to be in category 5.

DR. FRANCIS: Thank you, Dr. Davidson. Would the sponsors like to make any comment to this point?

MS. NORTON: We have no comments.

DR. FRANCIS: Thank you. Before we move to formal voting and Panel recommendations, I would like to give all

the Panel members an opportunity to ask any questions they may still have in their minds, and the sponsors one last opportunity to make any comments that they might wish the Panel to hear at this stage.

If the sponsors have no more questions, I would ask them to leave the table and take seats in the main body of the audience, please.

MS. CALVIN: Dr. Francis will now be calling for a motion and he will be asking the voting and temporary voting members to vote for whether this PMA supplement should be approved, approved with conditions, or not approved.

To reiterate. The voting members are Dr. George Birdsong and Dr. Diane Davey. Consultants appointed as temporary voting members for today are Dr. Machelles Allen, Dr. Juan Felix, Dr. Blake Nestok, Dr. Andrew Renshaw, Dr. Shailini Singh, and Dr. Robert Williams, Jr.

The Panel vote may take one of three forms:

1. Approval with no attached conditions;
2. Approvable with conditions; for example, resolution of clearly-identified deficiencies which have been cited by you or the FDA staff. These may include data clarifications or changes you would like to see in the draft labeling.
3. Not approvable. Section 515(d)2, paragraphs A

through E of the FDNC(?) Act, state that a PMA can be denied approval for any of five reasons, three of which are applicable to Panel deliberations.

The three reasons for recommending not approvable are:

1. There is a lack of a showing of reasonable assurance that such device is safe under the conditions of use prescribed, recommended, or suggested in the proposed labeling thereof;

2. There is a lack of a showing of reasonable assurance that the device is effective under the conditions of use prescribed, recommended, or suggested in the proposed labeling thereof; and

3. Based on a fair evaluation of all material facts, the proposed labeling is false or misleading in any particular.

To clarify the definition of safe, there is a reasonable assurance that a device is safe when it can be determined, based upon valid scientific evidence, that the probable benefits to health from use of the device for its intended uses and conditions of use, when accompanied by adequate directions and warnings against unsafe use, outweigh any probable risks.

To clarify the definition of effective, there is a

reasonable assurance that a device is effective when it can be determined, based upon valid scientific evidence, that in a significant portion of the target population, the use of the device for its intended uses and conditions of use, when accompanied by adequate directions for use, and warnings against unsafe use, will provide clinically-significant results.

If you vote not approvable, we ask that you identify the measures that you believe are necessary to place the PMA in an approvable form. Thank you.

DR. FRANCIS: You have heard the options before the Panel. Do we have a motion for one of those three options?

DR. ALLEN: I make a motion for approval, with labeling specifications. I believe that is choice number two.

DR. FRANCIS: Approvable with conditions.

DR. BIRDSONG: I second.

DR. FRANCIS: Second by Dr. Birdsong. Could I now have a show of hands to support the motion that this PMA be deemed approvable, subject to specified conditions?

DR. DAVEY: Do we have to specify the conditions now, or do we do it later?

DR. FRANCIS: I believe that is unanimous.

[There was a show of hands, and the motion was approved unanimously.]

I will now ask each of the voting members in turn to specify the conditions that you would like to see attached to this approval. We can start with Dr. Davey.

DR. DAVEY: Well, I will start. I will say, I am not sure I can remember everything right now, but if it is not brought up, if you could come back to me.

I think first of all, the labeling; we want to make sure that we handle the question of infections in high risk patients adequately, for both laboratories and clinicians to make clear as to how the instrument is used.

The second thing I want to say is that, in advertising, the manufacturer not be allowed to claim a new standard of care. So, those are the main things I wanted to say now.

DR. FRANCIS: Dr. Williams?

DR. WILLIAMS: The conditions I would like to see addressed are some type of clarification, reference primary screening, and some type of postmarketing analysis.

DR. FRANCIS: Dr. Renshaw?

DR. RENSHAW: The conditions I would like to see are, a more clearly-defined operation for what percentage of slides are actually being re-reviewed in the laboratory.

And a clarification of what the ranking system is doing.
Whether it is actually effective or not.

DR. FRANCIS: Would you just repeat the first part
of your recommendation?

DR. RENSHAW: More clearly defining how that 20 or
25 percent of slides that are designated No Review, what
that percentage is of. Is it the total laboratory slide
population, or only of those slides that are put on the
machine?

DR. FRANCIS: So, it is how that 20, 25 percent
designated as No Review are derived?

DR. RENSHAW: What population it is referring to.
Those that reviewed, or the total laboratory population.

DR. FRANCIS: Dr. Allen?

DR. ALLEN: Nothing to add. I guess just that I
feel strongly that populations that were not presented today
are noted.

DR. FRANCIS: Dr. Felix.

DR. FELIX: Agree.

DR. FRANCIS: Dr. Singh?

DR. SINGH: Just, again, is that high risk
populations were not studied.

DR. FRANCIS: And Dr. Birdsong.

DR. BIRDSONG: Clarifying the performance with

regard to infections, and particularly with herpes broken out, given the nature of that disease versus the other infectious diseases that are diagnosed.

Someone mentioned, the range of percentages and an average of percentages that you could expect to exclude instead of just, you know, up to 25 percent, because there was a significant range, and I think that should be in the labeling.

Finally, mention in the labeling, of explicitly stating there is an option to run high risk patients on the machine, because I think -- even though the study was not designed, doing that, I think there is enough data that has been presented to suggest that you are not going to harm anyone by doing that, and probably quite possibly bring some benefit to them, so.

DR. FRANCIS: Thank you.

DR. DAVEY: I guess -- Dr. Francis?

DR. FRANCIS: Dr. Williams, if you are asking for a formal postmarket surveillance -- oh, I am sorry, Dr. Birdsong -- you need to specify the --

DR. BIRDSONG: It was Dr. Williams.

DR. FRANCIS: Oh, it was you that specified -- I am sorry. You need to specify the actual parameters of postmarket surveillance. What it is that you want to see in

a postmarket analysis.

DR. WILLIAMS: Well, I guess it mainly pertains to the high risk population. I would like to see some type of study done on that.

DR. FRANCIS: Numbers of patients?

DR. WILLIAMS: On the high risk patients, the patients that were excluded.

DR. FRANCIS: I mean, do you have a recommendation for the number of patients that should be part of this postmarket surveillance?

DR. WILLIAMS: I did not have anything in mind.

DR. GUTMAN: I do not think it requires detail. We can take into consideration that recommendation.

DR. FRANCIS: Oh, okay. Dr. Davey?

DR. DAVEY: -- make a comment -- I am a little confused, I guess, about what Dr. Birdsong is saying in comparison to how it is going to be labeled, because on the one hand, if we are saying that if we need to explicitly say, it has not been done on high risk populations, then what Dr. Birdsong is saying is --

I mean, the only way I could see it as that, you would have to put in a comment that, you know, that the studies have not been done. The prospective study was not done in that. That there might be some value in using it as

a QC tool, but I would hate to not have those patients not get manually screened at all. I mean --

DR. GUTMAN: They are not going to be able to make a claim unless they have data to study it --

DR. DAVEY: Right. Right.

DR. GUTMAN: -- so, we would have to negotiate and take Dr. Birdsong's suggestion in context, but we would -- I assure the company, they will have very careful review of the labeling with our review team. We will try and carry out the spirit of what we have heard, but I doubt that they could start making claims.

They may be able to do research, or they may be able to convince you to do off-label use, but I do not think you are going to be able to convince us, based on the data they have in the submission, to start making claims for high risk patients.

DR. BIRDSONG: I maybe did not word my statement as tightly as I would like, because what you said is entirely consistent with what I am thinking. I do not mean to say the company should be able to just drop that statement from their labeling altogether, but from the way I read the statement, it would sound almost as if your -- I would not want it to be interpreted to mean that, you should not do that, under any circumstances.

As I said, you know, much earlier when I asked that question, you know, perhaps it is reasonable to do that and manually screen.

I just wanted to make it clear that that is not exclusively forbidden, and I guess, individuals can do what they want in their labs, but you know, I do think there is enough data to suggest there at least may be some benefit for high risk patients, and so, the label should not be written in such a way that labs run from doing that.

I think it should, as you said, add in a comment that you may want to do that in addition to manual screening or something.

DR. ALLEN: I think I would make two comments. That I think it is quite clear here among the Panel that we can only go with the data that has been presented to us here and the data is only with a particular population, or specifically excludes a high risk population. But, you are also hearing us say to you personally, that we think that, because of the very high sensitivity this tool could be used quite usefully in that very population, and perhaps you could reach out to us and we can help you get that data so you can in fact at some point change your labeling, appropriately.

The other thing I would ask the FDA to work with

the company for is to find out those specific things that cytotechnologists write in their comments -- multinucleate giant cells; tubular cells; endometrial cells -- so that the labeling can, you know, say that these things that were previously commented on cannot be commented on, because the technology -- or, maybe it can, but for those of us who are not cytotechnologists, but appreciate those comments and what the cells look like to the naked eye, can you address that or can you not? And can you put that in your labeling somehow, to help us to know what we are missing, or what we are getting?

DR. FRANCIS: Any other comments? Then what I am hearing is that the major issue expressed by the Panel is that the labeling should address the issues that the group of infections -- and the high risk group in particular -- were not -- did not form part of this study. And at the present time, the claims cannot be made for these particular groups.

The other issue, which I think came over reasonably strongly was the labeling needs to fairly clearly define the population from which the 20-25 percent No Review group are actually drawn. I see those as the major issues. Okay.

If there are no other comments, I thank the Panel

members for their deliberations and input. I thank the sponsors for their presentations, and now before adjournment, I believe Veronica Calvin has some closing remarks.

MS. CALVIN: I would also like to thank the Panel, and particularly, Dr. Francis for agreeing to serve as our Chairman today.

I would like to thank the public speakers for their input, and also the sponsor and FDA for their hard work, and thank you, Joan, for doing our transparencies. And also, before we leave, I would like to acknowledge Dr. George Birdsong. His term expires this February, and I just wanted to thank him for his outstanding service to the Panel, and his excellent contributions, and just so you know, we truly appreciate all of your input, and we look forward to calling upon you as a consultant in the future.

Lastly, the next meeting of the Hematology and Pathology Devices Panel has not been determined, however, the tentative dates for the remainder of the year are April 29 and 30, September 17 and 18, and December 10 and 11.

Thank you.

Agenda Item: Adjournment

DR. FRANCIS: This meeting is adjourned.

[Whereupon, at 4:20 p.m., the meeting was

concluded.]