UNITED STATES OF AMERICA

DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

+ + +

CENTER FOR DEVICES AND RADIOLOGICAL HEALTH

MEDICAL DEVICES ADVISORY COMMITTEE

+ + +

MICROBIOLOGY DEVICES PANEL

+ + +

March 8, 2019
8:00 a.m.

Hilton Washington DC North
620 Perry Parkway
Gaithersburg, MD 20877

PANEL MEMBERS:

BARBARA VAN DER POL, Ph.D., M.P.H.	Chair
RICHARD THOMSON, M.D.	Voting Member
LORI E. DODD, Ph.D.	Voting Member
BARBARA D. ALEXANDER, M.D., M.H.S.	Voting Member
LESLIE S. MASSAD, M.D.	Temporary Non-Voting Member
NICOLAS A. WENTZENSEN, M.D., Ph.D.	Temporary Non-Voting Member
ROBERT D. BURK, M.D.	Temporary Non-Voting Member
ELIZABETH R. UNGER, M.D.	Temporary Non-Voting Member
MONA SARAIYA, M.D.	Temporary Non-Voting Member
HERSCHEL W. LAWSON, M.D.	Temporary Non-Voting Member
WALTER KINNEY, M.D.	Temporary Non-Voting Member
REBECCA PERKINS, M.D., M.Sc.	Temporary Non-Voting Member
ANNA-BARBARA MOSCICKI, M.D.	Temporary Non-Voting Member
SARAH FELDMAN, M.D., M.P.H.	Temporary Non-Voting Member
TERESA DARRAGH, M.D.	Temporary Non-Voting Member
PATTI GRAVITT, Ph.D.	Temporary Non-Voting Member
BRADFORD M. SPRING, B.S.	Industry Representative
JOY McVEY HUGICK	Consumer Representative
ADEN S. ASEFA, M.P.H.	Designated Federal Officer

This transcript has not been edited or corrected, but appears as received from the commercial transcribing service. Accordingly, the Food and Drug Administration makes no representation as to its accuracy.

Free State Reporting, Inc.
1378 Cape St. Claire Road
Annapolis, MD 21409
(410) 974-0947

FDA REPRESENTATIVES:

UWE SCHERF, Ph.D.
Director, Division of Microbiology Devices
Office of In Vitro Diagnostics and Radiological Health

TAMARA FELDBLYUM, M.S., Ph.D.
Chief, Viral Respiratory and HPV Branch
Division of Microbiology Devices
Office of In Vitro Diagnostics and Radiological Health

MARY BARCUS, M.D.
Division of Microbiology Devices
Office of In Vitro Diagnostics and Radiological Health

STEPHANIE CACCOMO
Press Contact

FDA PRESENTER:

LUNA ZARITSKY, Ph.D.
Lead Reviewer
Division of Microbiology Devices
Office of In Vitro Diagnostics and Radiological Health

GUEST SPEAKER:

MARK SCHIFFMAN, M.D.
Senior Investigator
National Cancer Institute, NIH, DHHS

OPEN PUBLIC HEARING SPEAKERS:

JEFFREY ANDREWS, M.D., FRCSC
Worldwide Medical Director, Women's Health and Cancer
BD Diagnostic Systems
BD Life Sciences

CLEMENTINA COCUZZA, M.D., Ph.D.
Laboratory of Clinical Microbiology
Department of Medicine and Surgery
University of Milano-Bicocca, Italy

KEIVAN EITEFAGH, Ph.D.
CellSolutions, LLC

VARUNA SRINIVASAN, M.B.B.S., M.P.H.
National Center for Health Research

EDWARD EVANTASH, M.D.
Vice President, Global Medical Affairs
Hologic, Inc.

INDEX

INDEX

M E E T I N G

(8:04 a.m.)

DR. VAN DER POL: So I'd like to call this meeting of the Microbiology Devices Panel to order.

I'm Dr. Barbara Van Der Pol, the Chairperson of this Panel. I'm a professor at the University of Alabama at Birmingham in infectious diseases, and I am someone who's done a great deal of diagnostic evaluation studies in the field of sexually transmitted diseases diagnostics.

So I want to note for the record that the voting members of the Panel constitute a quorum as required by 21 C.F.R. Part 14. I'd also like to add that the Panel members participating in today's meeting have received training in FDA device law and regulations.

For today's agenda, the Panel will discuss and make recommendations regarding new or alternative approaches to the clinical study design and evaluation of devices detecting human papillomavirus (HPV) nucleic acid. These approaches will take into consideration scientific data generated since the approval of the first high-risk HPV screening device in 2003 as well as the effects of HPV vaccination on clinical studies of devices for HPV detection. Topics to be addressed at the meeting include clinical study design and comparator methods.

Additionally, the Committee will discuss potential changes to the HR HPV device indications for use, considering continually evolving cervical cancer screening guidelines. The Committee will provide expert feedback regarding the benefits and risks from the adoption of changes in each of the above topics and make recommendations for future HR HPV device evaluation, evaluation strategies that are both scientifically rigorous and least burdensome.

Before we begin, I would like to ask our distinguished Panel members and FDA staff

seated at this table to introduce themselves. Would you please state your name, your area of expertise, your position, and your affiliation.

DR. THOMSON: Good morning, my name is Richard Thomson. I go by Tom Thomson. I'm from NorthShore in Chicago, University of Chicago and Evanston Hospital, and my expertise is in laboratory testing.

DR. MASSAD: I'm Leslie Massad. I'm a gynecologic oncologist at Washington University in St. Louis and a colposcopist.

DR. ALEXANDER: Good morning, I'm Barbara Alexander. I'm at Duke University. I specialize in transplant infectious diseases, and I'm also a medical microbiologist.

DR. BURK: Robbie Burk, a professor at Albert Einstein College of Medicine, expertise in human papillomavirus genomics.

DR. UNGER: Elizabeth Unger from the Centers for Disease Control and Prevention. My background is as a pathologist, and currently, I work in public health of HPV.

DR. SARAIYA: Hi, I'm Mona Saraiya. I'm a public health physician by training, and my current expertise is in HPV and cervical cancer and public health.

DR. LAWSON: I'm Herschel Lawson. I'm retired from the U.S. Public Health Service and CDC, and I'm currently a professor of OB/GYN at Emory University in Atlanta.

DR. KINNEY: My name is Walter Kinney. I'm a gynecologic oncologist and a lapsed internist, and I am retired from the Kaiser Permanente system in Northern California.

MS. HUGICK: Good morning, my name is Joy McVey Hugick. I am the Consumer Representative today from Atlanta, Georgia, a former CDC employee working for the Division of Laboratory Sciences. So my expertise would be in laboratory sciences.

MR. SPRING: Yeah, I've got to hold it. I'm Brad Spring. I am the head of regulatory affairs for diagnostics at Becton, Dickinson and Company. I'm the Industry Representative on the Panel, and I think that's probably my expertise.

DR. FELDBLYUM: Good morning, I'm Tamara Feldblyum. I'm a branch chief in the Division of Microbiology Devices. My expertise is microbial genomics and infectious disease diagnostics, and the branch is responsible for respiratory virus diagnostics and sexually transmitted disease diagnostics, including HPV.

DR. SCHERF: Good morning, my name is Uwe Scherf. I'm the Division Director of the Division of Microbiology Devices and actively reviewing infectious disease diagnostics for 15 years.

DR. WENTZENSEN: Good morning, I'm Nicolas Wentzensen. I'm a cancer epidemiologist from the National Cancer Institute with a background in development, implementation, and validation of screening tools for cervical cancer.

DR. GRAVITT: Good morning, I'm Patti Gravitt. I'm from the University of Maryland School of Medicine. I'm a professor, and I am a molecular epidemiologist with an expertise in human papillomavirus and cervical cancer prevention through screening.

DR. DARRAGH: Good morning, I'm Teresa Darragh from the University of California, San Francisco, Professor of Pathology. I do cysto-pathology and cytopathology, and I also did colposcopy for 25 years.

DR. FELDMAN: I'm Sarah Feldman from the Brigham and Women's Hospital and Harvard Medical School in Boston, and I'm a GYN oncologist, and I exclusively take care of patients with HPV-related diseases of the large anal tract, cervix, vagina, vulva, and anus. Mostly cervix, though.

DR. MOSCICKI: And I'm Barbara Moscicki from the University of California at Los Angeles, and I'm an HPV molecular epidemiologist. I'm also a Professor of Pediatrics and chief of adolescent and young adult medicine and a colposcopist.

DR. PERKINS: I'm Rebecca Perkins. I'm a gynecologist and Associate Professor of OB/GYN at Boston University, and my expertise is in HPV and cervical cancer prevention.

I'm also a colposcopist.

DR. DODD: My name is Lori Dodd. I am a biostatistician at the National Institute of Allergy and Infectious Diseases. My expertise is in the statistical evaluation of diagnostic tests and clinical trials.

MS. ASEFA: I am Aden Asefa, and I'm the DFO, the Designated Federal Officer, for this meeting.

DR. VAN DER POL: Thank you.

For any of those of you who have not signed in with the sheets by the door, please be sure to do so before lunch, if possible.

And now I'm going to hand this over to Aden Asefa, the Designated Federal Officer for the Microbiology Devices Panel, and she'll make some introductory remarks.

MS. ASEFA: Good morning. I will now read the Conflict of Interest Statement.

The Food and Drug Administration is convening today's meeting of the Microbiology Devices Panel of the Medical Devices Advisory Committee under the authority of the Federal Advisory Committee Act of 1972. With the exception of the Industry Representative, all members and consultants of the Panel are special Government employees or regular Federal employees from other agencies and are subject to Federal conflict of interest laws and regulations.

The following information on the status of this Panel's compliance with Federal ethics and conflict of interest laws covered by, but not limited to, those found at 18 U.S.C. Section 208 are provided to participants in today's meeting and to the public.

FDA has determined that members and consultants of this Panel are in compliance with the Federal ethics and conflict of interest laws. Under 18 U.S.C. Section 208, Congress has authorized FDA to grant waivers to special Government employees and regular Federal employees who have financial conflicts when it is determined that the Agency's need of a

particular individual's services outweighs his or her potential financial conflict of interest.

Related to the discussions of today's meeting, members and consultants of this Panel who are special Government employees or regular Federal employees have been screened for potential financial conflicts of interest of their own as well as those imputed to them, including those of their spouses or minor children and, for purposes of 18 U.S.C. Section 208, their employers. These interests may include investments; consulting; expert witness testimony; contracts/grants/CRADAs; teaching/speaking/writing; patents and royalties; and primary employment.

For today's agenda, the Panel will discuss and make recommendations regarding new or alternative approaches to the clinical study design and evaluation of devices detecting human papillomavirus nucleic acid (HPV). The FDA is convening this meeting to seek expert opinion on the evaluation of risks and benefits from the adoption of changes.

Based on the agenda for today's meeting and all financial interests reported by the Panel members and consultants, no conflict of interest waivers have been issued in accordance with 18 U.S.C. Section 208.

Mr. Bradford Spring is serving as the Industry Representative, acting on behalf of all related industry. He is employed by Becton Dickinson.

For the record, the Agency notes that Dr. Mark Schiffman, who is an invited guest speaker with us today, has acknowledged his employer's receipt of testing products from firms at issue.

We would like to remind members and consultants that if the discussions involve any other products or firms not already on the agenda for which an FDA participant has a personal or imputed financial interest, the participants need to exclude themselves from such involvement and their exclusion will be noted for the record.

FDA encourages all other participants to advise the Panel of any financial relationships

that they might have with any firms at issue.

A copy of this statement will be available for review at the registration table during this meeting and will be included as a part of the official transcript.

For the duration of the Microbiology Devices Panel meeting on March 8th, 2019, Ms. Joy Hugick has been appointed to serve as Temporary Non-Voting Consumer Representative of the Panel. Ms. Natalie Portis, unfortunately, is not able to attend today, but she was being appointed as a Temporary Non-Voting Patient Representative. And for the record, Ms. Hugick serves as a consumer representative on the Gastrointestinal Drug Advisory Committee in the Center for Drug Evaluation and Research. These individuals are special Government employees who have undergone the customary conflict of interest review and have reviewed the material to be considered at this meeting.

The appointments were authorized by Russell Fortney, the Director of Advisory Committee Oversight and Management Staff in the Office of Special Medical Programs, on February 17th, 2019.

We would like to remind members and consultants that if the discussions involve any other products or firms not already on the agenda for which the FDA participant has a personal or imputed financial interest, the participants need to exclude themselves from such involvement, and the exclusion will be noted for the record.

FDA encourages all other participants to advise the Panel of any financial relationships that they may have with any firms at issue.

A copy of this statement will be available for review at the registration table during the meeting and will be included as part of the official transcript.

And before I turn the meeting back to the Chair, I would like to make a few general announcements.

Transcripts of today's meeting will be available from Free State Court Reporting.

Information on purchasing videos of today's meeting can be found at the table outside the meeting room.

The press contact for today's meeting is Stephanie Caccomo. If anyone desires to speak to the press, please speak with her or please see Ms. AnnMarie Williams at the desk outside the meeting room to obtain her contact information.

I would like to remind everyone that members of the public and the press are not permitted to the Panel area, which is the area beyond the speaker's podium. I request that the reporters please wait to speak to FDA officials until after the Panel meeting has concluded.

If you are presenting in the Open Public Hearing session today and have not previously provided an electronic copy of your slide presentation to FDA, please arrange to do so with Ms. AnnMarie Williams at the registration desk.

In order to help the transcriber identify who is speaking, please be sure to identify yourself each and every time you speak.

Finally, please silence your phones and other electronic devices at this time. Thank you.

Now I'll turn over the meeting to the Chair.

DR. VAN DER POL: We will now proceed to the FDA's introduction, which is going to be a little bit of a mixed methods approach where we will have Dr. Uwe Scherf approach the podium and give his presentation, and then Dr. Luna Zaritsky will also be participating in part of this presentation.

I will remind the public observers at this meeting that while this meeting is open for public observation, public attendees may not participate except at the specific request of the Panel Chair.

Dr. Scherf.

DR. SCHERF: Good morning. Again, my name is Uwe Scherf. I'm the Division Director in the Division of Microbiology Devices in the Office of In Vitro Diagnostics and Radiological Health.

First, I would like to welcome you all to Washington, D.C., the Panel members, the speakers, and the attendees of this CDRH Panel meeting. Second, I would also like to thank you, the speakers and the Panel members, for your contribution to public health and your service for public health here. Your contribution to this Panel is a clear public service, and FDA appreciates your recommendations and your feedback tremendously. Panel members, panel meetings are an important part of FDA's deliberation to resolve and address issues that are coming up with our review and our evaluations, and this panel meeting will help us to address these issues very much.

Please recognize, we will take your advice into consideration and really appreciate your feedback, and we are looking forward to all the comments you will make to us today, during the day and during the discussions. I would also like to emphasize that additional information can be provided for this meeting, and that can be submitted to the docket after the Panel meeting is over, and this can be done within 60 days of today's meeting.

So why are we here today? We want to discuss new approaches to the development and evaluation of devices intended for the detection of high-risk HPV devices, and we want to discuss whether these new approaches might allow for advances and innovation for HPV diagnostics and reduced burden to device manufacturers.

FDA has two missions. The first mission people are very much aware of, and that is to protect public health. We do this mostly by reviewing devices, in our case, before they go to the market in the premarket environment and by controlling them after they have been introduced to the market, the postmarket environment.

We have a second mission as well, and not everybody is familiar with that, and that

is to promote public health. We work very hard to accomplish the second mission as well, and today our meeting will be more focused on promote public health.

So we are contemplating about new approaches in the evaluation of high-risk HPV nucleic acid devices and how these approaches could benefit public health.

Why now? FDA, on its own initiation, has convened this Panel meeting as we did for the HPV test a year ago. We believe that now is the right time to reassess the way HPV diagnostic clinical trials are designed and how HPV clinical data is evaluated by FDA. And this is due to additional and accumulated knowledge in science and in the area of device safety and effectiveness and on vaccination and clinical trials.

Regarding the knowledge, the knowledge on the science, the study by Gage et al. indicated that HPV high-risk negative result is associated with less than 1% risk for developing CIN3+ over the course of 5 years. Also, the study by Tainio showed that within 2 years, about 50% of CIN2 lesions regress spontaneously.

Regarding the knowledge on device safety and effectiveness, clinical data is now available for five FDA-approved devices, and two out of these five approvals were based on very large clinical studies of 30,000 to 45,000 women. Clinical data is also available now on multiple collection devices and different liquid-based cytology media.

Knowledge on the vaccination and the trials. In 2017, the study by Oliver et al. reported that within 8 years of the vaccine introduction, the prevalence of vaccine-targeted HPV types decreased by 60 to 70%, and a very recent 2019 publication by McClung et al. reported, between 2008 and 2014, an 8% decrease in CIN2+ lesions caused by vaccine-targeted HPV types.

So the goals of today's meeting are we would like to have an open discussion and Panel recommendation of the outlined issue with current clinical study designs for HPV device evaluation. We would like to obtain the Panel's recommendation about the data

analysis approaches for performance evaluation of high-risk HPV devices. And we would like to hear discussion and Panel recommendation on FDA's proposals addressing other outlined issues presented by FDA and Dr. Zaritsky, which includes simplifying the intended use.

We have a tight agenda. In the morning, we have the presentations and the Open Public Hearing, and the afternoon will be occupied by questions to the Panel and the deliberation.

Lastly, I would like to sincerely thank the Panel Committee for their excellent work; from the Division of Microbiology Devices, Drs. Zaritsky, Feldblyum, Barcus, Bisht, Mielech, Conenello, and Gitterman, for help putting this Panel meeting together; from the CDRH Advisory Committee staff, Ms. Asefa and Mr. Swink; and from the National Cancer Institute, our invited speaker, Dr. Schiffman.

This concludes my introduction. Again, I would like to thank you for your public service. We really appreciate your input that we will obtain from you today, and I will turn the podium back to our Chair of our panel meeting, Dr. Van Der Pol.

Thank you.

DR. VAN DER POL: Thank you, Dr. Scherf.

And we'd like to now invite our guest speaker, and this is where we will have two people speaking. Dr. Mark Schiffman will speak, and Dr. Luna Zaritsky will speak intermittently. So we're also going to remind the public observers again that at this meeting, while the meeting is open for public observation, public attendees may not participate except at the specific request of the Panel Chair.

Dr. Schiffman.

DR. SCHIFFMAN: It's an honor to be here, to be invited for a talk of this length. My experience of many years, I've been studying HPV in the federal government since 1984 or

'83. If I read a talk for 45 minutes, there would be no one still awake because you're an expert panel, and in respect for that, I'm going to speak informally, and I'm going to go quickly, and if someone wants to slow me down or if I say something that you feel is not just a foundational truth, that's what I've been asked to do, is to try to put down what we know. And these are not opinions; these are supposed to be referenced and sort of noncontroversial foundations to your discussion later. So I will be talking about that foundation of knowledge that many of you helped create, and then I'm going to stop and the FDA will speak, and then I will speak again specifically to the knowledge that we have referent to each of those proposals or questions.

In terms of advancing, do I advance myself here? Yes. I have no personal conflicts of interest, but as mentioned, the National Cancer Institute does independent testing of screening test performance. So in that context, we have received HPV testing cytology results at reduced or no cost from many companies. We do consult pro bono, under supervision financially, for no conflict for anyone who asks our opinion.

The purpose of this presentation is to give background. It's limited to HPV of the cervix. It's limited to the topic of screening, so there won't be a mention of vaccination. It's relying on well-established referenced work. If I say something that you believe requires additional references, I would submit them to the record. Just let me know.

The parts of my talk will be to mention relevant aspects of HPV natural history and cervical carcinogenesis, evolving principles of cervical screening that are now under reconsideration first by several panels involved with screening itself, and the ASCCP is sponsoring a consensus conference for the management of screening HPV positive, other cervical screening positive results or abnormalities, of which the NCI is providing the scientific background for those revised guidelines. So I will be mentioning the general direction that those guidelines are taking as they are. They refer and inform some of the

decisions that are being considered. And then I'll give background relevant to each of the Panel questions after FDA has spoken.

The natural history of HPV and cervical carcinogenesis. I was parsing the title, "High-Risk Human Papillomavirus Nucleic Acid Detection Devices." What are high-risk human papillomaviruses? What's the difference between carcinogenic HPV as defined by IARC? I was the chair of the subcommittees that the last two times helped define what would be called a carcinogenic or definitely carcinogenic or the classes that are given to different certainties of carcinogenicity for HPV, and I was going to talk about the logic behind that distinction and how it's different than being high risk or for inclusion in HPV tests, so the necessity for inclusion in HPV tests.

And then I wanted to talk about high risk for what? I want to talk a lot about defining precancer, which is a vague term on purpose because precancer is a term we come up with to avoid using any of the common histopathologic nomenclatures because we define it just as the target of screening. A precancer is something that, if left undetected and untreated, has a high likelihood of invasion, and it turns out to be very important how that's defined operationally in terms of trial design.

I will not go over the molecular biology of human papillomaviruses. There are people on this Committee who know that much better than I do. We know it's a small DNA virus that is very, very slowly evolving, and I'm much more interested in how does that apply to diagnostics development.

So this is something that Dr. Burk developed. It is, I believe, the most current published chart of the evolution phylogeny of HPV, which is tightly linked to carcinogenicity. I don't really know if I have a pointer. I don't see that I do. But among the alpha papillomaviruses, HPV papillomaviruses, there's a clade, an evolutionary clade that is, for whatever purpose, linked to cervical carcinogenicity, and to highlight that alpha group

among the more than 200 HPVs, it's this group published by Dr. Burk and myself many years ago and it is the red group in the middle. Alpha 9, alpha 11, alpha 7, and alpha 5 and alpha 6 HPVs are unique in that evolutionary clade over time has developed the properties that, if persistent and in interaction with the cervical transformation zone, the zone of special susceptibility to cervical carcinogenesis, it's an actual ring of tissue that's called the T zone or the transformation zone, it has other names, right at the squamocolumnar junction. But that group, particularly those HPVs in alpha 9 that are HPV 16 related or alpha 7 HPV 18 related, are the ones that cause virtually all cervical cancer in the world, with much smaller contributions from alpha 7, alpha 5, alpha 6.

The point I'm trying to make is that there's an arbitrariness to our classification because each HPV type that I've just mentioned is actually hundreds or thousands of isolates or variants at the forensic or at the isolate level. We note that HPV 16 has variants that are -- these are subvariant level, they're A, B, C, D, that have evolutionary roots, and then within each, or the A, there's sublineages, A1, A2, A3, you know, D1, D2.

And then within that level there are isolates such that even in Oakland or Kaiser Permanente's study, you know, in the Bay Area of California and its surroundings, we found thousands of HPV 16s and multiple HPV 16s in a fair percentage of women carrying more than one isolate. We are going to be talking about HPV type. That is the level that we think about because currently we don't make use of variant lineage associations in screening, and there's no indication for those that's currently approved, but there are strong relationships of phenotype to genotype. Even within HPV 16 and others, there are certain types that are more risky, and they're related also to the co-evolution in the human host.

I won't be discussing that because it's not relevant to today's meeting. It should be noted, however, that the notion of type has an arbitrariness to it because we're stopping at that level rather than talking about individual, you know, isolates that would have to be

considered in terms of transmission if you're considering whether an exactly same infection had persisted or been transmitted.

How do we say that something is high risk or carcinogenic? Again, there's an arbitrary cutpoint. It was very carefully considered twice and in fact, admittedly, and I admit this again and again, we made a mistake the first time, which we retracted at the second IARC meeting, but we retracted it too late, and that error is still embedded in the DNA diagnostics that you'll be considering.

Because of the enormous relative risk of any particular HPV type as the necessary causal factor for cervical cancer, the relative risk for any type relative to being totally HPV negative is enormous, and therefore you can pretty much tell what the etiologic fraction of cervical cancer is for any given type strictly by how common is it alone, in other words, how often does it occur in invasive cervical cancers. Now, there are some that are co-infected, but cells themselves are not co-infected, but you can have multiple lesions on the cervix.

But in a cancer, it often has a single type that's driving the cancer, and in order of importance, you can determine the prevalence of HPV because it's not a hit-and-run virus, so it's persistent even in cell lines that are derived and maintained for decades, like HeLa or others. Basically, if you stop the HPV at that point where the part of the HPV that's integrated and persistent, the cell culture will die. So HPV, which is an extremely common infection, at some point becomes instead -- or parts of it, the E6 and E7 oncoproteins become growth factors and they drive the continued neoplasia.

So the change between infection, common infection, and a transforming oncogenic viral part is key to the understanding of what is infection to be not worried about that much because it's so common and so likely to clear under immune surveillance, be suppressed by immune -- subunit immunity and what is really worrisome because it represents one of the most powerful carcinogenic influences that we know of. That is a key distinction, that

switch, and it's at the molecular level, an interaction between the infection and the cell, but we have considered it for many years based on subjective morphologic visual criteria either at the level of looking at the cervix visually or microscopically with cytology or histology to correlate that with the sense of the underlying interaction between the virus and the host cell. Those subjective tests have done a tremendous job in reducing cervical cancer, including the Pap smear. However, there are correlations of the actual molecular process which we can now measure directly.

So what we can tell about invasive cervical cancer, for example, is that HPV 16 is qualitatively unique. It is the most important HPV. HPVs are numbered in order of their being added as unique types due to their L1 sequence, and so the numbers have no -- they're just historical in terms of it being in order. But as you know, around the late '70s, HPV 16 was found by methods we no longer use, by low stringency Southern blot, and then 18 was found after that in a cancer in Brazil, and 16 was a cancer in Germany, if I'm correct. And those turn out to be the most important types in the world because they, together, contribute a great deal of the fraction of cervical cancer.

After that, it becomes a little bit more regional in the types; 33, 45, 31, 58, 52 are certainly important types, but there is a slight variation in terms of where in the world they have different fractions. But, overall, those are always important types. Type 35 is unique in that it is uniquely risky to those of African ancestry.

After that, you get a group of HPVs that are found in very small percentage, as you see here, 1% or less of cervical cancers, but they are still considered sufficiently different in terms of relative risk and also generally what we know about their behavior in humans epidemiologically, but they're called carcinogenic: 59, 51, 56, 39.

Then we reach HPV 68, which is almost, if you look at it, 0.5% of cervical cancers. Where is the natural break here except for HPV 16? There's a break between 16 in terms of

etiologic fraction and the next group of like 18 through Type 35 or 52, somewhere around there, and then it goes down to ones that are nearly the same prevalence as non-carcinogenic HPVs, of which there are many that infect the cervix and can cause changes that look alike, that look like abnormalities that are often called ASC-US or LSIL or other kinds of minor abnormalities.

So we use HPV 6, which is the causal agent for condyloma acuminatum, as venereal warts or genital warts. It's found in 0.5% of cervical cancers, but we do not believe it to be a primary carcinogen except in the most rare circumstances. So the arbitrary cutpoint or the considered judgment was that if something's no more common than HPV 6, around 0.5, that it should be called, even if it's phylogenetically related to the carcinogenic HPVs. It could be put more in a class of possibly carcinogenic or 2B, and you can see that of those, Type 68 also is 0.5%, but it is ME-180, it is the driver of a cell line, and it's known to be the carcinogenic driver of that cell line. So it was put as 2A or probably carcinogenic.

But after that, we called things possibly carcinogenic in the rarest of circumstances, 1 in 20,000 cancers, 1 in 30,000 cancers, something that is vanishingly rare but could happen. And in that group were the -- importantly, were Types 66 and 53, which are pretty common types. They affect a lot of people. But we had, in a previous iteration of the committee, thought that HPV 66 was a probable carcinogen because it is linked to the development, prospectively, of carcinoma in situ or CIN3, which is the known surrogate for cervical cancer risk.

It turned out to be wrong. When we went from having 8,000 or 9,000 cancers in the world to 30,000, we found that HPV 66 dropped back and was not properly classified. So we took the unusual step of declassifying it. In the interim, many companies included HPV 66 in their assays, and by the nature of things in diagnostics, that remains because it takes a long time to change, and also, it takes a long time to change and have fewer types than someone else. So 66 is our personal mistake, mea culpa publicly, and we've been trying since then to get

new development to not include Type 66 because, in fact, it does not deserve to be included in terms of carcinogenicity.

There are many, many nomenclatures about how cervical cancer starts. The goal of screening is to find precancer. That's, again, this made-up term which we'll be using CIN3 as our best surrogate scientifically. But screening aims to find cancer before it develops; in other words, we are not screening for cancer. Although you can screen for cancer to downstage it in order to not have as much mortality, but screening in this field is different. In many different microbiologics, finding the agent implies disease. You find HIV, you find HCV, you find HPV, and you find something that's a pathogen, and that, in itself, means you're infected with something important. HPV is so common and so benign that mere detection of it is not akin to finding a pathogen or being a disease. It's an agent that overwhelmingly is -- and it's near ubiquitous now, and basically, it's a risk factor if it persists and if that change occurs between infection and precancer.

The peak of infection starts with the age of sexual intercourse of any given population, and it's enormously high. We only detect HPV in our point prevalence surveys if it's persisted to that day, so we don't even see the very, very rapid clears. But you can see, by this figure, that the HPV infection peak occurs very early, and then acquisition in all places goes down with changes with aging, of sexual behavior, and also acquired immunity, which is a partial acquired immunity to reinfection.

The second peak, which occurs some years later, it can occur only a couple of years later, the start of it, but then it peaks 5, 10 years later, is of detected or diagnosed precancer, which is defined in different ways that are very relevant to what we're talking about. Precancer is the precursor to invasive cancer, but that takes additional mutational events that break the contact inhibition and allow for invasion through the base of the membrane, and typically, there's a circumferential growth of the lesion for many years before invasion occurs, though

that's not invariable. It can be faster. So cancer arises as a slow process, many years after the first evidence of precancer if we were able to see it at its first clonal infection stage, but there's some time before we can detect it.

Now, that's represented here. These are based on real numbers. I'll be only talking about published literature except for the few times where we have, in preparation, very, very large studies from Kaiser Permanente in Northern California or other places that give us better estimates than we had before. And I'll say, in preparation, Dr. Demarco has been in charge of many of these studies.

But, basically, what this shows is the way we all think now about what happens if there's an infection. If we detect a point prevalent infection, which means this is not acquisition, this is point prevalent detection, and you see 100 percent of them very rapidly, very, very rapidly, most of them clear. What does clear mean? Clear is an ill-defined term, but it means mainly cell-mediated immune suppression by cells that are in surveillance in the stroma underlying the epithelium where the basal cells are infected with HPV.

So it does not mean total eradication out of the body for some infections. Some may be totally eradicated by AAV vector or some other mechanism. Some may be unfit and just disappear, but some are still in the body but in a state of perpetual latency or suppression, unless there is an immunosuppressive event like HIV immunosuppression, in which case they may reemerge under failed immune surveillance. But in most cases, in most places, and there are exceptions like in equatorial Africa, this curve would not be exactly the same, but in the United States, within a year or two, clearance is the rule. In other words, you cannot detect it even with sensitive DNA methods.

I think Dr. Burk called this pernicious papillomavirus if it persists, all of a sudden, a very, very common benign infection akin mainly to warty infections of the -- you know, that cause warts of the fingers or toes or face -- become something quite different, and after

years, persistence implies progression to precancer, which in this case, for these data, were represented as CIN3 such that by Year 7 you can see that there are virtually no infections left overtly that have neither cleared nor in a competing risk fashion progressed to CIN3. So long-term persistence of the most carcinogenic types almost implies progression. It's highly linked. Multiple years, for example, of HPV 16 is highly linked to having a precancer somewhere, even if it's still very small.

Then this flips on, and then there's a much longer time period that if there is CIN3 persistence, you gradually get invasion. Over the course of -- this is decades if you look at the x-axis. And this was, unfortunately, a result of a bad, bad idea of following precancers in New Zealand, documented by McCredie et al. and others, from what's called the unfortunate experiment. There's a lot of names for it. It's a disaster if someone had thought that they could follow a precancer without untoward effects, and people got invasive cancer, 30 to 50% depending on how you define it, over the course of their lives developed invasion. That's the only real knowledge we have directly, and fortunately, that's the only knowledge we have of observed invasion.

So it's just important to remember this basic fact. This is a 7-year follow-up of infections from Guanacaste, a full population cohort in Central America, that of 100 percent of infections, most cleared by 5 to 7 years and very rapidly in the beginning, such that a large percentage are gone within a year or two. And this is for point prevalent ones that have already persisted a little bit, which is important because the longer an infection persists, the longer it's likely to continue to persist. There are some that I'm sure didn't enter into this because they came and went so fast that we never even caught them.

The risk of progression to CIN3 differs by type. That's the fundamental important thing to remember about types. Progression risk varies slightly by viral load, at least for HPV 16 and the related alpha 9 types. But, in general, it's the type that matters. And I

would like to point out that what we know about the natural history is a result of mainly our knowledge of HPV 16 and how it produces squamous cell carcinoma of the cervix.

The much less common adenocarcinoma, or at least it used to be much less common before we had successfully attacked squamous in the United States, but adenocarcinoma has a different natural history and is linked to Type 18, 45, the alpha 7 types, a touch of other alpha 7 types like Type 39 and a subset of HPV 16. Not all HPV 16 variants are associated with adenocarcinoma. This is currently under study and is not as established as other things that I'm talking about.

An important fact, though, as if ignoring those that progress, all of the HPV types disappear with the same kind of a curve, this rapid, rapid benign clearance. And if you look at it, really, these curves are virtually in the same population of curve. It's a super-fast clearance initially, and then it slows down because the longer a type persists, the longer it persists even longer. So that sounds convoluted, but it's true that you see a leveling out of the risk or the clearance rate.

And so after around 3 years, 2, 3 years, you're starting to get into a worrisome situation where the likelihood of persistence and progression is starting to approach that of a chance of the virus going away. That's a really hard public health message to give people. If you wait, it will probably go away. But if it's still there in a couple years, then we need to worry about it. Most people do not like to hear that. They don't want to worry for years, so the natural way to manage HPV would be meticulously remeasuring as to whether it's still there. Virtually no one I know wants to do that because that involves worrying and waiting for a couple years. That's a real detriment to quality of life. So most people would like to know, at the time of identification, what's the chance that this is going to go away.

That's been the source of much of our research, is trying to find markers of likelihood of persistence and progression versus clearance at the time of detection. We call

that triage. It's not really a triage, which is a battlefield term for dividing people into thirds in terms of the importance of taking care of them, but we call it triage. And so triage tests are always essential for HPV because if you treat all HPV detected, you will be overtreating what's mainly a very benign infection. And yet, if you lose somebody who has an HPV that is going to go to cancer, a clinician is faced with a very bad situation of feeling they've failed the patient. So we're trying, within the world of HPV infection, to develop tools, tests that allow management, and that's the heart of the guidelines that we are doing now over a course of 2 years in collaboration with ASCCP and 20 other organizations.

The profound thing, though, is that absolute risk of CIN3 is type dependent. Sixteen is uniquely carcinogenic. This is follow-up from Dr. Kjaer's long-term follow-up study. Then there's that intermediate group that I discussed, and then there are the rest, which are very minimally risky but do occasionally progress to cancer.

Because of what I said about the difference between point prevalent first detection and persistent, if you just go to Kaiser Permanente and you look at the multi-year risk of CIN3 developing and you divide the infections even for 16 and 18 into finding it prevalently, you don't know how long it's been there and it could be an old infection, it could be just there yesterday. On the right, a point prevalent infection has a risk somewhere -- you know, for HPV 16, a 25.9% chance of developing CIN3, a very high risk. Eighteen, a little bit less.

But if an infection is new, meaning you knew they were negative and then they turn positive, the risk is much lower. This is going to be a recurrent and important theme in what we're talking about, in that a single negative with a good HPV test profoundly alters the risk of finding any biomarker that's used for screening in triage thereafter, such that if it's a new infection at any age, if the woman is 20, if the woman is 55, a new infection that's known to be new is much less risky because it's still at that first part of the curve that is

indicative of likely clearance. And that's true for all correlates of HPV, like an abnormal Pap smear, cytology, histology, anything that's related to HPV.

So high risk for what? I'm turning now to the next section. This has been very difficult to define. There's something called the heterogeneity of precancer that we've studied for decades. A surrogate endpoint is something that if you reduce it, you will reduce the cancer by the same proportion. That's a definition of a surrogate, is get rid of 20% of the surrogate and you get rid of 20% of the cancer. Formally proving something to be a true surrogate endpoint, like we use CIN2 or CIN3 as intraepithelial precursors, is not ever validated because we don't follow them and prove that that relationship holds, except for that one experiment. So we are guessing at surrogacy, and for clinical use, we use CIN2+, and I'm going to be discussing how that can really mess up the consideration of DNA diagnostics. Or RNA diagnostics, sorry. Nucleic acid diagnostics.

When I first started, I had an artist adapt this very well-known figure that had been around for years by Ferenczi and others, and basically it just shows that there's lots of terms that have to do with the continuum from HPV infection, normal epithelium, all the way to invasive cancer, and those terms were studied rigorously by Mark Sherman, who arrived at this, what he thought was a more correct set of terms, which I blurred out on purpose because they're entirely blurry conceptually.

(Laughter.)

DR. SCHIFFMAN: As Dr. Sherman would agree, given that he left cervical pathology and is now working on other things -- that's a joke, for the record -- that all the systems we have, from the Papanicolaou to dysplasia to CIN to the Bethesda system, are all approximations with a high degree of trouble of reproducibility and accuracy in definition. Particularly, I want to address the issue of CIN or any intraepithelial neoplasia scale.

The CIN scale was developed by Richart and Barron, and I remember this happening.

Ralph Richart did a study in which they had a lead-in period of several months for people who had CIN1, what he called CIN1, which was called mild dysplasia. So he made sure that the lesion was already persistent before he even called it CIN1. He then saw that almost all of them progressed through increasingly severe disease and decided to turn it into a tiered three-part CIN1, 2, 3. He could have said CIN1, 2, 3, 4, or 5 or whatever. He chose three. There was no known -- I have no understanding of any scientific evidence that there's actually a three-part natural history progression from CIN1 to CIN 2 to CIN3. I don't believe those entities are biological entities. They're morphologic measurements. And Dr. Richart himself, in 1990, says that two would be better than three. And, in fact, that's true because we have HPV infection and its correlates, and then we have precancer. We transformed transforming infection, and it is, in fact, a two-part natural history, but there's a three-part clinical scale.

That was corrected because CIN2 doesn't mean anything. CIN2 has no biological natural history state, and when we did the ALTS trial, and this is published in *JAMA* by Mark Stoler, if you look among the CIN2/CIN2 agreements, see the 70 in the middle, of things that the consensus biopsy quality control committee thought were CIN2, less than half were called CIN2 by the community. I've seen studies in which our panel, nothing has been called CIN2 by everybody. CIN2 is an equivocal term between infection and precancer.

So we've been trying to divide it based on biomarkers and to eliminate it, and that has been done by the LAST criteria, and it's important to do something like that because CIN2, being equivocal, is not a great surrogate of cancer risk. Look at the HPV 16 etiologic fraction in this huge series of typed -- these are cases from our Kaiser population called the PATH cohort, and if you look at 16, which is supposed to be -- is predominant in cancers, you see that the first types in cancers are 16, 18, 31, 45. Just like everywhere. Those are very important types. But 16 in CIN3 is over half, but in CIN2 it's only 36.3%. CIN2 is caused

by all kinds of types that do not go to cancer. That's a really important point. CIN2 is a mixture; it's a construct that's really between infection and precancer, and because of that, it is not a good surrogate for cancer, because if you reduce CIN2 by a certain percentage, you would not reduce cancer by that same percentage because a fair proportion of CIN2 is caused by other types.

So true precancer is an unknown subset of CIN2/CIN3 adenocarcinoma in situ, which is the glandular equivalent. P16 is our most used biomarker, but it's not perfect. There's these other things that I'm not recommending or anything; they are things that are being researched, and I just put them down to mention that there are other biomarkers that indicate a higher risk of something being precancer and infection rather than just an infection.

And with that, I'm going to turn to some of the basics of what's happening to cervical screening, and then I'm going to go back to FDA.

So the same diagram. The goal of screening is to find precancer, true precancer, while ignoring the very, very common HPV infections that can just go away if you don't do anything. So primary prevention is, of course, happening through vaccination. That's already impinging on what we're seeing in 21- to 24-year-olds and will continue to impact screening. That was not one of the questions I was asked to address.

Important historical precedents and newest developments. We're now emphasizing the actual risk of the molecular process of the positive predictive value, negative predictive value, and risk stratification from knowing the HPV status and its likely age and how long it's been there. There's much less emphasis now on sensitivity for CIN2 or CIN3 because that's known. You're basically looking for how much HPV is present in a referent standard that itself is an alloyed referent standard. It's a correlative of the infection. The main process is a molecular process that has to do with the HPV interaction with the tissue. So we're

replacing morphologic subjective terms in the new guidelines with much more precise and accurate HPV-related terms.

Everything has to do with the 2 by 2 table. Everybody needs to just refer constantly back to the 2 by 2 table and its extensions. Instead of looking at columns, in other words sensitivity being over a + c, of a detection by a nucleic acid diagnostic, though that's very important in all the derived statistics of sensitivity and specificity like ACRSC curves and we still use that extensively, we're increasingly, in terms of management guidelines, looking at the rows. Like if a woman is positive, what does that mean and what should you do? If a woman is negative, what does that mean? And those are much more related to prior probability and post-test probabilities, and that's where the field is going in terms of risk-based guidelines.

As HPV supersedes cytology for these reasons, things that were dependent on using cytology alone, like ASC-US, triage becomes obsolete. I've spent more time studying ASC-US than about anyone in the world, and ASC-US means I don't know what this is. The majority of Pap smear abnormalities or non-normal Pap smears are in the "I don't know what this is." ASC-US, ASCH, AGC, all of them mean we're not sure if this is normal or not, and that's by far the majority of non-normal cytology. So HPV is much more reproducible, and it's also more predictive.

It's unclear whether co-testing, adding the two together, cytology and HPV, is cost effective. I would actually say it's not unclear, but that's an opinion because really only high-grade cytologies are important in terms of risk stratification, and they are rare and if a woman is HPV negative.

There's many pending developments in triage tests. There's this continued desire to get better and better triage tests among HPV positive women. It's very difficult to anticipate all the options and the developments. It's so rapid it's going to be very hard to

regulate in a constant way that lasts.

Basically, we're going to a risk-based approach where we just -- we figure out the risk somebody has, either minimal, low, medium, or high, of development of CIN3, which we take as our best proxy, and then depending on that, that we have cutoffs that are being determined by the medical societies representing American medicine, and that is being put into a risk matrix through clinical trials, medical record data, high-quality observational prospective studies.

We decide the risk over time, now up through 5 years, and based on risk, there's risk action thresholds that are being devised by the clinical groups, and that consensus of more than 20 organizations is going to lead by the end of this year, starting in June, then with a public comment period lasting until October. ASCCP will be hosting the vote on the endorsement of the consensus guidelines, which will be put into a look-up table application for clinicians in the United States to deal with all of the different tests that are now on the market on the basis of risk.

I'm going to stop now and turn it back to FDA and then for the specific questions.

DR. VAN DER POL: So Dr. Luna Zaritsky from the FDA will now give her comments. We'll hold questions until the end.

DR. ZARITSKY: Good morning, everybody. We would definitely like to thank you all for being here today. This is an exciting day as we discuss new approaches in the evaluation for high-risk HPV nucleic acid detection devices.

So what I'm going to do is shift gears a little bit. Dr. Schiffman provided a very comprehensive scientific and biological background, so what I'm going to do in this talk is discuss more the topics that the FDA has been facing challenges with over these past few years that we'd like to focus the Panel discussions on for today. And then I'm going to hand it back to Dr. Schiffman, who's going to present some more data that specifically addresses

some of the questions that we're asking today.

So our overall goal here at the FDA in the near future is to start issuing new standardized recommendations to industry when they're seeking FDA approval of high-risk HPV devices. So the purpose of this meeting today is threefold.

First, we'd like to discuss the scientific topics that will help inform FDA recommendations on device development and evaluation.

Second, we'd like to obtain Panel recommendations for approaches to clinical study design and the establishment of performance characteristics.

And, thirdly, we'd like to explore potential pathways for innovation in HPV device regulation that are less burdensome and more streamlined, because the more efficient the approval process is, the more timely access the community will have to these devices.

So I'll just provide a brief background on HPV devices. We just heard an extremely comprehensive scientific background. I'll just focus on the device background.

So as Dr. Schiffman has already said, HPV devices are used in cervical cancer screening to assess a woman's risk for harboring an early treatable precancer because the point is to identify these women early and treat them before they're too far along in the disease process. We're defining precancer as a lesion that could progress to cancer if left untreated.

So there are over 200 HPV types, 40 of which infect the genital tract. However, as Dr. Schiffman has explained earlier, the vast majority of these infections will clear on their own within the first few years. So we don't want these devices to detect every single infection that's out there. We only want them to detect the infections that have the highest risk of causing cancer, so the clinically relevant infections, where we're defining clinically relevant as types with known carcinogenicity, so the high-risk types.

So listed here are the 14 types that HPV devices detect; 66 is in italics, and

Dr. Schiffman already provided a background about 66, so I won't repeat that. But even with types that are carcinogenic, so even with these high-risk types, not all of them will progress to cancer or will cause cancer. So the point of the devices, again, it's not to detect all of those high-risk types, but only those with viral loads that are high enough to be associated with some sort of clinical manifestation. And the clinical manifestation that we have traditionally been using is CIN2+ lesions on the cervix.

So because for HPV devices the threshold that distinguishes between a positive and negative result is based on a clinical endpoint, the devices are said to have clinical cutoffs or clinical cutpoints. And because there's no known viral load for all of the types that's associated with the CIN2+ lesions, the clinical cutoffs for HPV devices are set separately by each manufacturer and validated in a clinical study using CIN2+ as an endpoint.

So high-risk HPV devices are Class III medical devices that are regulated by the Division of Microbiology Devices in the Office of In Vitro Diagnostics and Radiological Health at CDRH. So this means that they require a PMA, or premarket approval, prior to marketing.

So outlined here is just a brief regulatory history of HPV devices. So in 1991, the first HPV device was approved to detect high- and low-risk HPV.

And then 9 years later in 2000, the indication was revised to include ASC-US triage for women to determine the need for colposcopy referral, so this is HPV testing in women with ASC-US cytology to determine the need for referral to colposcopy.

And then in 2003, HPV testing was approved to be used in routine cervical cancer screening as an adjunct to cytology in women aged 30 years and older. So this is what we today refer to as co-testing.

And then in 2014, the indication for primary HPV screening for women aged 25 years and older was approved, where the HPV test is being used as the initial screen.

So the next two slides just show the wording of these indications. So for the ASC-US triage, the indication states that in women aged 21 years and older with ASC-US cytology, the device is used to determine the need for referral to colposcopy.

The adjunct indication states that in women aged 30 years and older, the device can be used together with cytology to adjunctively screen to detect the high-risk HPV types.

And, finally, the primary screening indications states that in women 25 years and older, the device can be used as a first-line primary screen where women who test negative should be followed up in accordance with physician's assessment. Women who test positive for HPV genotypes 16 and 18 should be referred to colposcopy. And women who test HPV positive but 16/18 negative, so 12 other positive, should be evaluated by cervical cytology to determine the need for colposcopy referral.

So the next question is why are we here? Why do we want to reevaluate our current strategies? And there are several reasons for this.

First, we have a broader knowledge of cervical carcinogenesis from published research. It's been 16 years since HPV testing was approved for routine screening, and there were a lot of unknowns back then that aren't unknown anymore, especially pertaining to HPV and its role as a biomarker for cervical cancer. So we want to see if we can use what has been discovered on the research side to help inform our regulatory decisions so we don't end up reinventing the wheel every time we get a new submission.

The second reason has to do with HPV vaccination and the decreased prevalence of vaccine-targeted HPV infections. We've already seen this with 16 and 18, and now that we have the 9-valent vaccine, we do expect to see this in the future. So this is something that is definitely going to affect our clinical study design because the pool of the positives that we need to evaluate these devices is slowly going to be shrinking.

The third reason has to do with evolving screening and patient management

guidelines. So this is something that is not only going to affect the standard of care of the women enrolled in the studies, but it's also going to affect our indications for use.

And lastly is safety and effectiveness data from previous HPV device approvals. So there are five HPV devices that are currently approved, and we not only have the safety and effectiveness data from the clinical studies that were submitted to the FDA that supported approval, but for several of these tests there is data on how they perform postmarket in a real-world setting. And this data comes from clinical research as well as real-world evidence.

So we're armed with a much more comprehensive understanding of how these tests perform and the risks and the benefits associated with these tests, especially the risks associated with a negative result. So we want to see if we can leverage some of this information to make the evaluation of HPV devices more relevant to what the current benefit-risk profile of these devices are today, as opposed to what they were over a decade ago when some of these studies were originally designed.

So here is a brief outline of the topics that we'd like to discuss today with the Panel, or that we'd like the Panel to discuss today. So what I'm going to do for the remainder of my talk, I'll introduce each of the topics that we'd like to discuss and briefly describe some of the challenges that FDA has been facing when dealing with these topics. Then I'll describe some potential proposals that may address some of those challenges. However, if anybody here has any other ideas or approaches that may work that have not been mentioned, we definitely encourage active discussion of those. And then, lastly, I'll discuss some of the benefits and risks associated with some of these proposals. But, again, if anybody here today can shed some light on additional benefits and risks or even some risk mitigation strategies, that's something that we are definitely excited to hear about.

So something that I will mention is that some of these proposals that we're talking

about today are actually things that FDA has done in certain submissions as different situations arose, and they worked out pretty successfully. So now we want to see if we can open up the discussion into something a little broader to see what approaches we can adopt as something more standardized across the board for all new devices that are coming to market or that are coming to the FDA.

Okay, so the first topic that we'd like to discuss today has to do with the clinical study design. So the clinical studies supporting HPV device approval are very large. These studies have enrolled anywhere between 10,000 and 40,000 women, depending on the indications sought by the sponsor. And these large sample sizes are primarily due to the low prevalence of CIN2+ and CIN3+ in the U.S. screening population.

So just as an example of this challenge, in one of our recent approvals for the adjunct indication, so this is women with normal cytology aged 30 years and above, out of a population of over 8,000 women, there were only 7 with CIN3. And this challenge is expected to become exacerbated as HPV vaccination rates increase because this will likely lower the prevalence of the vaccine-targeted types as well as the lesions that are caused by these vaccine-targeted types. So if we continue to keep doing what we're doing, the study sizes are just going to end up progressively getting larger and larger.

So the first discussion topic today is how can sponsors reasonably obtain a sufficient number of high-risk HPV positive and CIN3 positive women in a clinical study supporting high-risk HPV device approval?

So we have three proposals to start off the discussion, but again, if anybody here has additional proposals that we haven't mentioned, we please welcome discussion of that. So these proposals could potentially be used by themselves, or a mixture of the proposals could be used depending on the particular submission.

So for the first proposal, a sponsor would conduct a prospective study. They would

enroll, say, X number of women. They get what they get in terms of HPV positivity and CIN3+. But then they could supplement that prospective population with subjects recruited from colposcopy clinics, where the prevalence of HPV infection and associated disease is expected to be higher.

Now, the second proposal would be to allow the use of archived specimens from file banks, repositories, laboratories, etc. Now, these archived specimens could potentially be prospectively collected but then archived in some biobank, for example, or they could potentially be retrospective specimens that are used to supplement a prospective study. And from our interactions with sponsors, it does appear that these types of specimens are available.

Now, the third proposal would be to cap the proportion of vaccinated women in the clinical study. So this is something that we have allowed in the past. Say the clinical study can cap the vaccination rate at 10% and the remaining 90% of women could be non-vaccinated. However, we did want to bring this up for discussion considering that as vaccination rates are increasing over time, the intended use population for these devices will become a vaccinated population.

So there are benefits and risks associated with these proposals. For Proposal 1, supplementing with women from colposcopy clinics, the benefits are that there would be a higher prevalence of HPV positive women and the sponsor would be more likely to enroll women with CIN3+.

However, the risks that this population would not be representative of the screening population. These would be women who have had abnormal screening results or abnormal screening histories and may be farther along in the disease process.

Another risk is that these women may have higher viral loads. Again, if they're farther along in the disease process, they may have higher viral loads, so specimens from

those subjects may be easier to detect than the specimens that are coming from a screening population.

Now, for the second proposal, utilizing archived specimens, the benefit is that the sponsor could test a large number of specimens without the burden of prospective enrollment. And, additionally, depending on the institution where they get these specimens, there's potential access to longitudinal information, so follow-up data in the patient records. So we can essentially do a longitudinal study without actually having to wait whatever number of years.

However, the risks are that there is potentially limited resources, and also the colposcopy/biopsy procedure that was done that led to the histology diagnosis in the patient record would not have been performed under a standardized protocol. It would be whatever the standard of care was for that patient, so that could introduce variability.

Now, the third proposal to cap the vaccinated population, the benefit is that obviously with a primarily non-vaccinated population, there will likely be a higher prevalence of HPV positivity and CIN3+, provided that there's limited herd immunity.

Now, the risk, though, is that the performance may not necessarily be representative of the device once vaccination rates are higher and the distribution of the circulating genotypes will be different.

Another risk is that this may not even be effective once herd immunity plays a greater role and the non-vaccinated population starts to look more like the vaccinated population.

So the first question that we're posing to the Panel today is would the Panel recommend one or a combination of the following three proposals to increase the number of women positive for CIN3+ and/or high-risk HPV in clinical studies:

1. Supplementing from referral clinics

2. Utilizing archived specimens
3. Capping the vaccinated population

Again, if you have any other proposals that would work better, that would be great for the discussion.

Okay, the second topic that we'd like to discuss, or we'd like the Panel to discuss today has to do with the colposcopy referral protocol in clinical studies supporting HPV device approval. So as we mentioned earlier, HPV devices are evaluated against a clinical endpoint, so that means colposcopy/biopsy, histology adjudication, that's all part of the clinical protocol. And I do want to point out that in these studies, referral to colposcopy does include a biopsy removal, either a directed biopsy if lesions are present or random biopsies, and in some cases, endocervical curettages are taken as well.

So for women involved in these clinical studies, they all have three screening results. They have a cytology result and they have two HPV results, one from the investigational device that's under evaluation and the second from the FDA-approved device that's used as standard of care. So any woman who was positive for HPV by either device is referred to colposcopy. Additionally, any woman with abnormal cytology, so with ASC-US or greater, is also referred to colposcopy. In addition, women with normal cytology, so women who are NILM who are negative by both devices -- or, I'm sorry, a random subset of the women who are NILM and negative by both devices are also referred to colposcopy. And that random subset is around 5% or so.

So the two populations that we want the Panel to discuss, that we're requesting the Panel to discuss today, are the NILM double negatives, so negative by both devices, and the ASC-US double negatives. And this is primarily due to what the standard of care is for these women right now.

So over the past decade or so, I'm sure, as many of you already know, the screening

and management guidelines have evolved due to data that illustrates the high negative predictive value of HPV testing and ruling out CIN3+, as well as the high regression rates of non-high risk HPV-related cervical abnormalities.

So the 2012 ASCCP consensus guidelines state that women who are NILM/HPV negative should return for rescreening in 5 years, and women who are ASC-US/HPV negative should return for rescreening in 3 years. So this means that the subset of the NILM double negative women and all of the ASC-US double negative women participating in a clinical study for approval will undergo a procedure, which is colposcopy/biopsy, that they would not normally undergo as part of standard of care, and subsequently, they will be exposed to the risks associated with these procedures that they wouldn't be exposed to had they not enrolled in a study.

So the next discussion topic is, given how guidelines have evolved over the years, is it still appropriate to have these populations of women undergo colposcopy/biopsy for the evaluation of a high-risk HPV device?

So there are benefits and risks to colposcopy referral for these populations. The benefits are that this would ensure an unbiased clinical performance evaluation because all populations with different combinations of HPV and cytology results will have histological diagnoses for evaluations and we'd be able to adjust for verification bias.

Now, the second benefit is that this would prevent making assumptions regarding the CIN2+ and CIN3+ state in double negative women, so in women negative by both devices. So if these women were not referred to colposcopy, we would have to make an assumption that they all would have negative biopsies. And we do know that for NILM/HPV negative women, the 5-year CIN3 risk is about 1 in 1,000, and for ASC-US/HPV negative women, the 5-year CIN3 risk is about 5 in 1,000.

Thirdly, the third benefit is that we'd also get a better estimate of CIN2+ and CIN3+

prevalence in women missed by multiple HPV devices.

Now, there are also some risks. So the first risk associated with referring these populations is the identification and overtreatment of transient lesions that would have regressed on their own had the woman just followed standard of care and didn't come back until another 3 to 5 years later, but because she was immediately sent to colposcopy, the lesion was treated. And studies have suggested that overtreatment of cervical lesions could be associated with deleterious reproductive outcomes.

Now, the second risk is the potential underestimation of device sensitivity. So we do know that there are some lesions that have non-high risk HPV etiologies, so these could be lesions that are caused by low-risk HPV infections or even inflammation. And these lesions have a high probability of clearing on their own and are thus not a part of the causal pathway to cervical cancer and not what the HPV devices are really designed to detect. So including them in your endpoint for performance evaluation may underestimate device sensitivity.

And the third risk is the potential optimization of devices to cross-react with low-risk types in order to have better sensitivity. So this is something that's especially pertinent with the CIN2 lesion. Dr. Schiffman already provided some really nice scientific data showing how lesions could be present -- or sorry, how different HPV types could be present in a CIN2 lesion but don't actually progress to cancer. So we don't want to encourage a situation where sponsors may want to optimize their devices to cross-react with these genotypes in order to have a better sensitivity estimate.

So the second question that we're posing to the Panel today is, for these two populations of women, do the benefits of colposcopy referral for the assessment of verification bias outweigh the risks associated with the procedure and potential overtreatment? And we do ask that you discuss this for the two populations separately

because they do have different risks associated with them.

Okay, so the third topic that we'd like the Panel to discuss today has to do with the indications for use for the HPV devices. So as I mentioned earlier, there are three main indications for use for which high-risk HPV devices have received FDA approval. There is ASC-US triage for women aged 21 years and older, there is adjunct to cytology for women aged 30 years and older, and there's primary screening for women aged 25 years and older.

However, we are facing some challenges with these indications. The first challenge is that tying device indications to specific populations that are outlined by the clinical guidelines at the time, as well as specific clinical actions such as colposcopy referral based on device results, may end up in misalignment with future changes to clinical guidelines, and we may end up in a situation where older devices are indicated for uses that are no longer recommended by the clinical community anymore.

So just as one example of this, we can take the ASC-US triage indication which states that women aged 21 years and older with ASC-US cytology should be triaged to colposcopy based on an HPV result. Now, this is something that was consistent with the clinical guidelines at one point; however, now women who are aged 21 to 24 with ASC-US cytology are no longer triaged to colposcopy based on a positive HPV result. So this is something that not only results in an indication that's not quite aligned with what's going on right now, but it also makes the studies harder for the sponsors because when they're conducting their clinical studies, there just aren't women in that age group being tested for HPV anymore.

Now, the second challenge is that certain populations overlap between the indications resulting in redundant analyses. So for PMAs, the approval of an indication is based on the data that the sponsor submits supporting that particular indication. So there's essentially a data package associated with each indication that a sponsor is seeking. However, we can see with these indications that the adjunct population, which is women

aged 30 years and older, is also part of the primary screening indication, which is women aged 25 years and older. So this results in some data analyses essentially being submitted multiple times, which doesn't seem to be a very efficient use of time and also doesn't align with the FDA's least burdensome principles.

So the next discussion topic is how do we simplify device indications for use and make them independent of potential changes in clinical practice in the future?

So one proposal would be to just modify the indications for use, and this modification would consist of three parts. First, we could consolidate the indications to encompass one general screening population. So instead of there being three different indications, each representing a different population that's being tested, we could consolidate them into one indication for the general screening population. So an example of such wording could be that this assay should be used to test women presenting for routine cervical cancer screening to assess the risk for cervical dysplasia and cancer.

So I do want to point out that even if the language in the indications for use statement is for a general population, the FDA will still intend to evaluate device performance for different subpopulations within that larger population, so different age groups, cytology groups, and any other population that the Panel thinks would be necessary to evaluate. That could still happen, and that data would still be in the labeling for people to see.

So the next part of the proposal would be to remove references to specific triage tests in the indications for use statement. So as we talked about earlier, for primary screening it states that women who are HPV 12 other positive could be tested by cervical cytology to determine the need for a colposcopy referral. So, as of now, cytology is the only approved triage test, but we do know that in the not-so-distant future there's going to be a plethora of additional triage tests that are going to be coming down the pike. And if we

continue to specifically state what triage test should be run after an HPV positive test in the HPV manufacturer's indications for use, then we may end up with manufacturers continually wanting to update their indications for use to keep up with these new technologies. But every time that there's an update to the indications for use, that's another PMA supplement that's submitted to the FDA. So, this is something that could create an administrative and regulatory burden, considering that the device that's actually the subject of the PMA supplement would not have changed at all.

So, the third part of this proposal would be to remove references to specific clinical actions based on results. So, instead of saying, "If you get this result, go to colposcopy," the language would be more general to accommodate any sort of future changes in the guidelines.

So, taking Parts 2 and 3 together, the indications for use statement could say something along the lines of: "Women who test positive or negative for HPV types should be triaged and followed up in accordance with professional guidelines, the physician's assessment of screening and medical history, and other risk factors."

Now, again, there are benefits and risks to modifying the indications for use statement like this. The benefits are that the enrolled population would be relevant to current screening practices, and this would accommodate future changes to screening populations. For example, if in the future, vaccination results in a change in the age [at] which screening begins, then having this kind of generalized wording would still work with that. We wouldn't have an indication for a population that's inconsistent with the populations that are being screened.

The next benefit is that new triage strategies will not require a change -- sorry, a change to the device indications for use, and subsequently would not require another PMA supplement.

And, lastly, this would prevent misalignment with future changes to clinical guidelines. For example, if in the future, the screening algorithms that dictate clinical actions become more complex, where the clinical actions are not just based on the results of one or two tests, but you have a multitude of other factors that are being plugged into this risk algorithm, such as vaccination status, medical history, age, then having this type of generalized language would eliminate the situation of having an indication that states a certain population should be managed in a way that could be inconsistent with how they should be managed when using this more complex risk algorithm.

Now, the risks are that this generalized wording will not provide specifics on the populations who are tested in the study. In order to see the populations that were included, you kind of have to dig through the labeling to see those analyses.

And another risk is that the device would not be analyzed as part of a screening algorithm. So, the data won't show a direct depiction of the device performance when it's used in a specific way or with specific triage tests.

So, the third question that we're posing to the Panel today is: Regarding the indications for use, do the benefits outweigh the risks for:

A. Consolidating the indications to encompass one general screening population; and

B. Removing references to specific triage tests and clinical actions

We'd also request that you please discuss any potential risk mitigation measures that the FDA should adopt if a new IFU statement were to be used. For example, are there specific populations that should be included in every study? Are there any sorts of warnings or limitations that should be included in the labeling?

So, the fourth topic today pertains to the data analyses that the sponsor should conduct in order to support the indications for use. Now, there are five HPV devices that

have been FDA approved, and as I talked about before, several of these devices have been clinically validated in a real-world setting post-approval using clinical research and real-world evidence.

In addition, as I'm sure we all can agree here, that the usage of HPV as a biomarker for cervical cancer is firmly established science. So, taking all of this together, it begs the question of whether we're at the point right now where there are enough clinically validated tests on the market with sufficient safety and effectiveness data so that we can start doing device-to-device comparisons.

So, the next discussion topic is how can we begin to incorporate device-to-device comparisons for performance evaluation in premarket submissions?

So, we'll discuss two potential proposals for doing this, and these two proposals are very different in that one proposal is a paradigm shift from what we're doing today in that it will not include a clinical endpoint comparator, so it would not include histology, whereas the second proposal is more consistent with what we're doing today in that it does include histology or a clinical endpoint comparator.

So, the first proposal, this is the one that would not include a clinical endpoint, it would be to establish performance against a composite comparator. So, this composite comparator would be composed of three FDA-approved tests that have already been clinically validated against a clinical endpoint, and then the majority result from these tests would be the comparator result. And then the performance of the investigational device would be assessed against this composite comparator.

Now, I do want to point out that this is a simplified illustrative example of this type of comparator. We only talk about overall positivity and overall negativity on this slide; however, we do appreciate that with devices that have different genotyping outputs, this would be much more complex than just what's illustrated here and that there would have

to be additional studies to validate the clinical cutpoint for those outputs. So, we completely appreciate that, but for the purposes of the discussion today, we want to focus on more of the big picture question, which is the general acceptability of utilizing a comparator that does not include histology. So, we'd like to focus just on that big picture question as opposed to what the details of the comparator should be if you have devices with different genotyping outputs.

So, as I said before, that this comparator would not include histology, so we would not be assessing the risks associated with CIN2+ and CIN3+ associated with different device results. Colposcopy referral and histology adjudication would not be part of the clinical protocol, and all patients would be managed as per standard of care.

Now, the second proposal is a shift from what the first one is in that it would still include a clinical endpoint in the evaluation, it would still include histology. But in this one, we'd be assessing the relative performance between an investigational test and an FDA-approved test against a clinical endpoint. So, performance parameters such as sensitivity against histology or a clinical endpoint that includes histology would be assessed for the investigational test and the FDA-approved test, and then the ratios of these estimates would be evaluated to assess non-inferiority of the investigational test.

So, again, there are benefits and risks to each of these comparators. So, for the composite comparator, the one without a clinical endpoint, the benefits are that this may lead to more efficient clinical studies because colposcopy/biopsy, histology adjudication, none of that would actually be involved in the study.

Another benefit is that multiple devices could be assessed in a single patient, so that would provide some useful information as to how the devices stack up against each other.

And, thirdly, because histology would not be a part of the evaluation, we wouldn't run into the scenario that we talked about before where certain populations of women

would be referred to colposcopy against what their standard of care is.

However, there are risks. So, the first risk is that this type of evaluation could only be done for the common result outputs between devices, so overall HPV positivity and then whatever the common result outputs are between the devices. If there is a device that has a unique genotyping output, then additional analyses would have to be conducted to validate those channels.

The second risk is that we would not be able to assess the clinical relevance of different device results. We could have devices that have different results, but without histology, we won't know if that difference is actually clinically relevant or not.

And the third risk is that we would not have histology to confirm that the population enrolled in the study includes women at highest risk.

Now, for the second proposal, when assessing relative performance, the benefits are that this would provide information about new and approved devices and the clinical relevance between those differences. So, if there are differences in device results, we would have the data to know if these differences are occurring in a woman with versus without disease.

The second benefit would be that it would add objectivity to the performance evaluation. So, we do know that histology can be subjective, or at least more subjective, than molecular tests. So, having the results from an FDA-approved test in the evaluation could potentially help normalize some of the variability associated with purely morphological assessments.

Now, the risks again, that this could only be performed for common outputs between devices, and because histology would be included in the evaluation, it still may need a large study sample size unless we can think of some newer approaches to make the clinical study size a little bit more -- or a little bit less burdensome.

So, the fourth question that we're proposing to the Panel today is to please discuss whether the following types of data evaluations are acceptable for the assessment of safety and effectiveness for new high-risk HPV devices: Adoption of a molecular composite comparator method, so adoption of a method that does not include a clinical endpoint, versus evaluation of relative performance between a new and approved device against a clinical endpoint.

So, we'd ask that you please discuss the benefits and risks to both of these proposals, as well as any other proposals that the Panel may have, as well as minimum acceptable performance criteria. So, when answering this question or when discussing this question, if you believe that the evaluation should include a clinical endpoint, then that will bring us to our final discussion topic today, which would be the clinical endpoint comparator.

So, as we discussed earlier, HPV devices are evaluated against a clinical endpoint, so this is CIN2+/CIN3+. However, there are challenges with utilizing a purely clinical endpoint comparator. First, and this is something that we actually have seen in the past, is that lesions may have non-high-risk HPV etiologies, which may have less validity as a surrogate endpoint for cervical cancer.

Now, the next two challenges are not so much things that we've seen in the past, but these are things that we may anticipate seeing in the future as HPV vaccination rates are increased.

So, with 9-valent HPV vaccine, we expect to see a decrease in prevalence of the seven high-risk types that are targeted by the vaccine, as well as a decrease in prevalence of the lesions that are caused by these vaccine-targeted types. So that means, out of the remaining lesions that are left, those that have non-high-risk HPV etiologies may represent a larger proportion of what's left compared to the proportion they represented in the pre-

vaccine era when you still had the lesions that were caused by the vaccine-targeted types in the mix. And this is something that could potentially affect the performance estimates of newer devices if we're just using a purely histological endpoint.

Secondly, the distribution of HPV genotypes that are causing the lesions in a study will change depending on the proportion of vaccinated and non-vaccinated women who are enrolled in the study, as well as herd immunity effects in the populations where these women are enrolled from. So, depending on a device's ability to detect certain genotypes, performance may look different in different studies that have different make-ups of the populations.

So, the final discussion topic is, "What clinical endpoint comparator should be used to assess performance?"

So, one potential proposal could be to utilize a mixed histological and molecular comparator that consists of both histology and HPV typing. So, the benefits of this type of comparator are that this would inform the type of false negative results a device yields. If a device has a negative result, we'd be able to see if it's in a specimen that truly does not have high-risk HPV nucleic acid versus one that does. We'd be able to assess the clinical performance for genotyping outputs that are unique to the investigational device, and we'd more accurately be able to assess a device's ability to detect lesions that are caused by vaccine-targeted and non-targeted types.

However, the main risk is that it may be difficult to determine what histology/genotype combination constitutes a positive or a negative comparator result because ultimately, we don't know which lesions are going to progress to cancer and which ones are going to resolve on their own. At best, we would only have a risk, and then the question would be, "Where do you draw that line where the risk is high enough to be a comparator positive and low enough to be a comparator negative?"

So, this table here outlines the different histology diagnoses and HPV genotype combinations that could potentially occur in a clinical study. So, on the left-hand side we see the range of histology, and on top is the HPV result where it's divided into the types that are HPV -- sorry, the specimens -- that are HPV negative, low-risk HPV positive, and high-risk HPV positive, where the high-risk HPV positives are further subdivided into the types that are targeted and non-targeted by the HPV vaccine, or the 9-valent HPV vaccine. However, if the Panel would like to discuss the genotypes separately or in any other combination, that's fine, too.

So, if a mixed comparator is something that is within the realm of possibility to do, the question would be, "What would the comparator result be for the different combinations?

So, we started filling in this table for some of the more obvious results. For example, for histological diagnosis of CIN3+, that is 16 positive, that would obviously be a comparator positive result. Conversely, for a negative biopsy in a woman who's HPV negative, that would obviously be a negative result. However, there are a bunch of empty cells here for some of the combinations that are not so black and white. So, again, if this is a comparator that the Panel thinks is acceptable to be used, we will request your help in interpreting this table to determine what the comparator result would be for these different combinations.

So, the final question today is, "If the Panel recommends assessing HPV device performance against a clinical endpoint comparator, is utilizing a mixed histological/molecular comparator acceptable, and if so, how should the combination of HPV result and histological diagnosis factor in when assigning comparator positive and negative results?"

So, in summary, I just want to show you the questions that we're going to pose to you today, and I will point out that all of these questions will be read again to you this

afternoon, and you will have several hours for deliberation, but I'll just very briefly summarize each of these questions again.

So, the first question pertains to the clinical study design and potential approaches to make the study size more efficient.

The second question pertains to colposcopy referral protocols and whether the benefits outweigh the risks for colposcopy referral for the ASC-US double negative and NILM double negative population.

The third question is for the indications for use statements, whether the benefits outweigh the risks for consolidating the indications and removing references to specific triage tests and clinical actions.

Our fourth question is whether a comparator that uses a clinical endpoint versus one that doesn't use a clinical endpoint, the risks and benefits associated with those.

And, finally, if the Panel recommends assessing performance against a clinical endpoint comparator, whether the mixed histological/molecular comparator is acceptable and how should the combinations of results factor in when assigning comparator positives and negatives.

So, in closing, I would just like to thank everybody at the FDA, from our division, from our office, and from our Center, who played a role this past year in trying to develop this complex roadmap for the future of HPV device regulation, and I especially want to give a huge thank you to everybody in the Division of Microbiology Devices who played a role in contributing their ideas and discussion.

So, with that, I will hand the microphone back over to Dr. Schiffman, who will present some data that's specifically pertinent to some of the questions that we're asking today. So, thank you.

Oh, and then after that, we will have a Q and A session with both of us.

DR. SCHIFFMAN: I was at the 1991 approval of the first HPV detection device, and I think we've come a long way towards complexity and understanding. So, I'm glad I'm on this side rather than your side because these are tough questions, but I'm not going to make any other normative opinions. I'm just going to very, very quickly get to the Q and A session by saying that we have some specifics for these and they all refer again to the 2 by 2 table.

So, would you recommend any of these different ways of getting more precancer, which is the total a + c on this table, how to do it? In our current studies, NCI recognizing that only 0.5 to, at most, 1% of any screening population will have what you're looking for at the screen. That means that you've got to screen, you know, basically tens of thousands of people if you're going to derive your cases entirely from your screening population. If only prevalent risk is important, we are using case-control or case-cohort designs.

The length of negative predictive value, the reassurance limits, use of these approaches because, as mentioned, the HPV positivity proportion increases as you approach diagnosis because the lesion is getting bigger and more obvious for CIN3. And we saw this in this complex looking thing that I'll show you that was published in *JNCI*. And if you look at CIN3 on the right-hand side in the middle, red is -- this is all the cancers that were found in Kaiser Permanente Northern California, or CIN3s. This is the CIN3 panel, this one, and it's sensible if you can just follow the logic of it, but it's a little scary to look at.

So, it's going backwards in time from the right to the left, from time of diagnosis of CIN3s. There are many, many thousands of CIN3s diagnosed between the year 2003 and the current day, and this was published, and the question is was the woman -- if she had a co-test, because at Kaiser they co-test every 3 years -- was she HPV positive, was she cytology ASC-US or greater?

And going backwards, you see this red, which is positive by both cytology and HPV

dropping off. Orange -- or whatever that color is, brown, orange -- is positive by HPV alone, the little green is positive by cytology alone, and blue is negative by both, missed by both. And you see some years before, 2, 3 years before, you start to see a drop-off in positivity.

So, what does this mean? If you're doing a screening study and you're supplementing with colposcopy clinic CIN3s, you will be looking at bigger cases, bigger lesions in your supplement group than you'll find in your screening group, which will be looking in advance, and that affects the length of time that you can claim that the negative HPV test reassures against CIN3 because the test may very well capture everything that's just about to be diagnosed histologically, like in colpo clinic because it's already abnormal. But does it lead to the kind of 5-year or more reassurance, a negative predictive value for those tiny lesions that you're trying to pick up in a screening kind of setting? And, of course, we're looking for extended intervals now in our guidelines. Is that at all clear? I don't know whether --

The next one, archived specimens. We have lots and lots of archived specimens, a million specimens or something from our archive. I would very much like to turn that into a profit center for the NCI. That's a joke, too, which I apologize for.

(Laughter.)

DR. SCHIFFMAN: And we are approached often for archived specimens. The whole point is proof of comparability. We store the specimens, and you can reiterate a study from those stored specimens, and we do that for novel devices that we're studying. Those archived specimens are gold, they really are great, though most of the cohort specimens you collect sit there and nothing -- the woman goes along and nothing happens, so you're collecting a million specimens of things that do not relate to the disease and you store them at great expense for years. They do exist, those archives. They are fundamentally valid because they're deep frozen aliquots and they're stable for DNA. We have less experience

working with RNA. The question is how do you prove that they're comparable to what would've happened in a new collection? I'm just going to raise that issue.

Vaccination population capping. Because of her protection, that is really -- among women 21 to 24 at KPNC, you're seeing a generalized decrease in outcome. I don't know how feasible that is in the United States based on the experience that was published by Castle et al. in *Preventive Medicine* last year.

Question 2, this is an important question because verification bias adjustment is trying to look for how many cases are missed among the HPV test negatives. So, basically, you're looking for what is the drop-off in sensitivity, incomplete sensitivity, and also what is the impact on negative predictive value. And, often, something on the order -- in a study of 30- to 40,000 people, a thousand double negatives or whatever are invited and maybe 800 go and you find extremely few because of, as I said, the Gage paper, which showed if you look at years since baseline and then you look at the middle panel for CIN3+, it shows that over 5 years, if someone is HPV negative, and this is with hybrid capture, Qiagen hybrid capture, which is not the most sensitive assay but it is used for this purpose, so it certainly holds for the others as well, for the other DNA tests at least. You can see that it's 1 in 1,000 in 5 years and very few right away that will have CIN3 if they're HPV negative.

And so if you refer 800 people and you have a 1 in 1,000 5-year risk, your predictive CIN3 is something on the order of one or zero, and then you need to multiply by the sampling fraction of 30 or 40, and it totally blows up that cell and overwhelms the estimate of sensitivity such that you can get results in indications that are frankly wrong. You will find a lot of CIN2s, and they're often of types that were not meant to be targeted even by the HPV test under consideration.

And so, the attempt to correct very slight verification bias can result in amplified misclassification bias. I call it amplified misclassification bias because of the sampling

fraction, it amplifies this sort of finding, the difference between zero and one or two odd cases that are missed. I can give specific examples, but I think we're supposed to avoid specific named products or whatever. But there are indications and labeling that I believe to be false, demonstrably false. So, HPV negative ASC-US has a very low risk, this was already just mentioned, and HPV negative NILM has extremely low risk, less than 1 in 1,000.

The idea of generating one screening population and just using the entire screening population to get to the other traditional indications, I don't have much to say about that. I would rather -- I'll speak more on -- I spoke already about ASC-US and that it's a fairly equivocal and obsolete term if we move away from primary cytology alone, which is a movement in the United States, but the important thing is this issue of going back, removing references to specific triage tests and clinical actions.

In one of the recent approvals there was something called a candidate strategy that was indicated; what do you do if it's positive, what do you if it's negative? This is problematic going forward as the FDA -- as Luna said, basically because we have so many different triage possibilities coming forth. Even more so as published by Castle et al. just recently in *JNCI*, we're going to be moving to a country that is using HPV testing. A single negative HPV test is much more impactful than a single negative cytology.

So, if you look at the risk of all of these common terms that are used as benchmarks for different actions, like HSIL or LSIL, ASC-US or NILM or whatever, for positive and negative, you see that after one negative the country is altered. So, let's just say that we make guidelines and then 90 or 95% of people who are screened with an HPV test are negative at any one time, the 30-plus. So, coming back the second time within a few years, and which you do and your recommendations are meant to last for more than a few years, but when people come back for their second screen, most have been negative the last time and all of those findings will be new or incident.

And as I mentioned, the risk after -- of an incident finding is so much lower than a prevalent finding, such that even in HSIL, even that's HPV positive, goes from risk of 52 for 3 years to 33, LSIL goes from 6 to 3. The lower you go, the less it's impacted. But most of the abnormal results don't mean the same thing, and the ASC-US triage tests don't mean the same thing following a negative, and almost anything you can mention related to HPV, if it's incident doesn't mean the same thing. So that continues the more HPV negatives that accrued. And so, making a management decision as part of the indication, an indication, as we've talked about for HPV 66, stands for a long time, how are we going to continually improve the indications and labeling? And so, for the guidelines, we're continually changing them through a process of continual review, but I don't know how the FDA will do that. I'd like to say that everything I'm saying is meant to be a background. I don't have the answers to any of these. I'm just giving background.

In terms of 4 and 5, which I put together, I'd like to point out that HPV tests do have -- the DNA tests do have about an equal cutpoint, they have a sweet spot where they hit, so this is a predictors study by Cuzick et al., but they're not all the same. The RNA test is more specific. In my view, it's slightly less sensitive. We do not have long-term prospective data on the RNA test to know whether it's exactly the same. However, if you had to choose multiple predicates, you'll realize that it's not that they're all exactly the same, so what are you averaging in terms of calling it a composite?

I'm not even going to get into suggestions. I think that it's not my place to suggest, so I'm going to just leave it to you, and I'm going to thank you for the opportunity to assist in this, what I consider, very important update, and I'm really sincerely looking forward to your conversation and see what you have to say about these important issues.

Thank you very much.

DR. VAN DER POL: Thanks to both Dr. Schiffman and Dr. Zaritsky.

Now, at this time, we'll take clarifying questions from the Panel. So these are just brief questions to our two presenters, if there are any points of information or fact that they'd like to have to clarified, so this is not quite the discussion portion yet.

DR. UNGER: Thank you both for the presentations, and this question is for Mark, and it has to do with what you called the complicated picture. I think it's Slide 39. And you were emphasizing that a limitation of using people referred to colposcopy was that the lesions would be larger than you'd identify in a screening population. But if you required the women -- so women go to colposcopy for a variety of reasons, but if you required them to have been enrolled in a screening program prior to having an abnormality, wouldn't that mitigate that impact somewhat?

DR. SCHIFFMAN: Absolutely. In our studies, we make sure that the colposcopy clinic is the colposcopy clinic that's of that screening population in that catchment area with proper epidemiological principles so that if someone from the screening clinic is abnormal, they would go to that clinic. All you're doing, really, is taking 100 percent sample of everybody who goes to colposcopy from that catchment area, and you're doing a sampling of the people who are screened. So it's a valid design. I'm just saying that in our experience, you're catching a lot -- you're also catching some people who have more overt lesions.

DR. UNGER: Because of the 5-year lag, but again, some people would come into screening having not been in a regular program, and they would be at much more risk of having the large lesions. So, if you excluded those and required them to have been under regular screening, I think that would be very analogous to the size of lesions that you would detect by screening.

DR. SCHIFFMAN: I agree. There's a bimodal distribution of cases, particularly for cancers of, you know, of ones that are symptom detected and screen detected. They can

be much different in age, and they can be distinguished by asking for history, and you can get exactly the efficiency that you're talking about.

DR. VAN DER POL: Thank you. That was Dr. Unger.

And Dr. Wentzensen has a question now, so --

DR. WENTZENSEN: Yes, I have one for Mark and one for Luna, if that's okay.

So, Mark, can you specify your presentation on the risk of persistent infections? So, you clearly showed that, one, your persistence increases the risk a lot, but I think that there's kind of some confusion about very long-term persistence because the risk in these women actually is much lower because you take out all the disease that happened early on. And so, if somebody had the same type for 5 years, that risk is lower; is that correct?

DR. SCHIFFMAN: It depends on the type. HPV 16 in the PATH cohort is still getting all cases. HPV 51, if it's a long-term persister, there are -- there's a very, very small subset of people who just have carriage. It's really a small group. So, you're right that there is a surge in detection of disease, and then for the lower-risk ones, it's true what you're saying. 18 and 45, though, are linked to poorly detectable, you know, glandular lesions. So long-term persistence of those, in my experience, is worrisome even if it's very long. But it depends on which type. And so, it again indicates that you have to think of this as a viral type-specific or viral type group-specific phenomenon.

DR. WENTZENSEN: Then I have one quick question. Today, are we talking about any HPV tests, including RNA? So, we consider RNA?

DR. ZARITSKY: Yeah. So, if you think that there are some things that should be distinguished between a DNA and RNA test that you can't generalize across both, then we definitely would like you to bring that up.

DR. WENTZENSEN: One more clarification regarding the genotyping of endpoints. You are referring to cytology-based genotyping to really -- if you're talking about combining

a CIN3 endpoint with a molecular classification, that is based on -- it is not tissue-based genotyping, or are you including that in the consideration?

DR. ZARITSKY: So, the original thought was to be on the liquid-based cytology sample, but if you have any other thoughts about genotyping on the tissue, that would be something that's definitely relevant to the discussion.

DR. MASSAD: I was going to ask that, too. Do we have any data about tissue-based genotyping?

DR. ZARITSKY: We do not at FDA. All of our assays are run on the liquid-based cytology sample.

DR. VAN DER POL: Dr. Feldman had a question.

DR. FELDMAN: Hi, thank you. This is a question for Luna. So just when you were talking about the history of cytology approvals, you talked about the type of like brush or whatever. So, in terms of the HPV approvals, if we're going to broaden these approvals, I wanted to know are we going to consider, you know, self-administered vaginal swabs? Are we going to consider how the specimen is collected? And I also saw online this week some new thing that you can send away for and do your HPV test, and so, yeah, that's online this week. So, I wanted to know whether that would be part of what we would want to be considering as to how it was acquired, how frequently it was approved for that sort of thing.

DR. ZARITSKY: So, the discussion today will -- we'd like to base it on clinician-collected specimens. Self-collection is something that's a different conversation, but just to keep this focused on the clinician collected.

DR. VAN DER POL: Dr. Moscicki.

DR. MOSCICKI: This question is for Luna and perhaps another one for a combination. But on your point number four, data analysis support indication and the one that was for, including histology, I didn't really understand the benefit when you say it adds objectivity.

Either we believe histology or we do not believe histology, but having these tests, I don't see how it's going to make us not believe it or believe it, especially if you're going to include that as a composite.

And then just maybe the second question for both, that we really haven't discussed using persistence, type-specific persistence, as an outcome in any of these, especially if we think it's so important and as well as using not CIN2.

DR. ZARITSKY: Okay. So just to answer your first question, so that point was based on our experience in the past where when we'd have a handful of CIN3s or something, occasionally we would get one that just seems like an anomalous specimen where it's negative by multiple FDA-approved tests, it's negative by PCR sequencing, so something is kind of off with that specimen in terms of whether it actually does include nucleic acid or not. But because we only have a handful of those specimens, you know, it will have a large effect on the estimates for sensitivity when just using the purely histological endpoint.

So, having additional molecular tests can kind of help us figure out what's going on with the specimen. Is this a problem with the investigational test, is it just not picking up what it should be picking, or is there something going on with the specimen, maybe collection or something like that? But we definitely would like to hear your discussion on those topics.

And then the second question about persistence. So, our tests, we don't measure persistence. It's just we evaluate whether the test is detecting an infection at that time, but we don't have indications for persistence at this point.

I don't know if, Mark, you want to add?

DR. SCHIFFMAN: Even in ALTS, when we had distinguished panel consensus, histopathology for CIN3, I think eight of the cases were negative by everything. And so even in terms of reliability and reproducibility of CIN3, there are some cases that are ambiguous

in terms of is this a look-alike or is this neoplasia. So, the question is if all of the FDA-approved tests say that that's negative for high-risk HPV, could that composite be compelling in terms of not calling a new negative by the new test, a false negative or at least putting in a special category for estimation of sensitivity?

DR. VAN DER POL: Dr. Massad has a question.

DR. MASSAD: Thanks. I had two questions for Dr. Zaritsky.

One is if we use a general indication for use that, in your wording, calls on manufacturers' devices to be used according to professional guidelines, can they still be marketed if there's no professional guideline?

DR. ZARITSKY: Maybe Dr. Feldblyum could have an answer to that.

DR. FELDBLYUM: Well, this is just thinking on this slide here. We assume that they're always going to need professional guidelines, but if there aren't, certainly the language also adds that according to medical practice and physician's assessments. So, we would leave medical practice to medical practice rather than FDA regulating how medical practice should be done.

DR. MASSAD: So, the answer is yes?

DR. ZARITSKY: Yes.

DR. VAN DER POL: Dr. Dodd.

DR. DODD: Yeah, I have a question for Dr. Schiffman. So, the question is how much of our knowledge and assumptions about the type carcinogenicity depends on the viral population dynamics prior to the introduction of HPV vaccination? So, if the vaccination campaign alters the distribution of type, what's the potential for a shift in those fundamental assumptions in terms of carcinogenicity?

DR. SCHIFFMAN: I worried about that for a long time. Could there be some kind of niche replacement or a possible -- if we get rid of HPV 16 as the bad actor, will some related

type evolve to take its place? We've exhaustively studied that. I'd be glad to share all the references. It appears, to the best of our knowledge studying co-infections, which are very common, and the fate of each of them in the presence of each other, that they do not crosstalk, they do not synergize, they do not antagonize that we can find. The cervix is huge compared to the HPV and that little clone of cells. So, we do not see any niche replacement. Plus, these are extremely slowly evolving DNA double-stranded viruses that evolve pretty much by drift at the pace of the human genome itself. So, it's not a rapidly responsive niche or evolutionary pressure kind of a system, which makes it much more of a good target for public health intervention because there's not escape evolution that we've ever seen. And so, our knowledge has been based mainly on pre-vaccination, but we have now over 10-year experience in post-vaccination populations, and we do not see anything.

Now, of course, the percentage among infections of the other types goes up as you take out the very prevalent 16 or whatever, so it does change the composition and the riskiness of the residual pool; however, that's just a matter of averaging basically. And so, I could go on and on because -- but I won't. But, basically, there's not synergy, and there's not antagonism, and they don't seem to crosstalk. So, I think that you can remove one. They're each running their own race side by side. And so then basically you've gotten rid of a bad one, you drop risk, and I don't see any compensation happening.

DR. DODD: So just to clarify, then, the role of persistence of these other types, you're saying, is not -- you know, in 10 years from now, you don't expect to see that resulting in a spike or an increase of carcinogenic lesions due to the changing distribution of these other types?

DR. SCHIFFMAN: In a very poorly screened environment, the very slow and less risky types, because you haven't treated the more aggressive earlier 16 types that evidence themselves more aggressively and overtly. You could get more like HPV 39s if you're poorly

screening everybody and it's not included in the vaccine and you just let that go because they weren't treated, the T zone wasn't removed. This is all extremely theoretical, but it would be that 39 is coming along at its own very slow risk. I can show you individual type risks out to 9 years now, and some of these types, like 39, that are marginally carcinogenic, it's flat. If you don't find something in the beginning, you find almost no risk, and the cumulative risk curve just goes like this and then goes flat. So, it's a risk of fundamentally zero after the original point prevalent detection.

DR. VAN DER POL: Dr. Darragh.

DR. DARRAGH: Thank you.

I actually have a question for each of you. Luna, when you talked about the composite comparator without histology, you suggest three FDA-approved tests. Since you also say that, you know, the labeling is by medium, by sampling device, why three rather than two, because some FDA-approved tests are only approved under two mediums, excuse me, or under -- so say for SurePath, you can use cobas or Onclarity, but you wouldn't have a third comparator there.

DR. ZARITSKY: Yes. So that's definitely something that we would have to work out some of the details on. For some analytes, when there aren't FDA-approved tests available, we do work out some validation strategy for the purpose of the test. However, for the purposes of the discussion today, that level of detail will be addressed once we have some sort of bigger picture. But, yeah, that is an extremely valid point, and these are things that we at FDA kind of have to think about and to establish the details of those types of protocols, for sure.

DR. VAN DER POL: Dr. Perkins.

DR. DARRAGH: Dr. Schiffman, a little bit more of a philosophical question. If we, in fact, have good uptake of the 9-valent vaccine, I should say when we do, in fact, have good

uptake of the 9-valent vaccine, will the number of cervical cancers attributable to the other high-risk HPV types rise to the level, public health level, that we should be screening for them?

DR. SCHIFFMAN: I think that's beyond my purview today. I really am trying to restrict myself to what we know to be fact. So, conjecturally, you could guess as well as I could. So, I don't know, that's a value -- I don't want to talk beyond what I'm told to do. I'm on a tight leash.

(Laughter.)

DR. VAN DER POL: Dr. Perkins, I think this is the last question.

DR. PERKINS: I have two very, very specific questions. One is if you're utilizing archived specimens, who owns those specimens, and would there be a risk of them being used up by industry and, therefore, not being available for other uses? That is the first question.

DR. ZARITSKY: Yes, that is a risk. Maybe Dr. Schiffman might have -- since you have so many archived specimens, you could share some -- shed some light onto that.

DR. SCHIFFMAN: Fortunately, you can do a lot of tests off of our residual aliquots of liquid-based collections, but we try to be responsible stewards. We set up usually either a committee for some kind of use criteria. We believe that there is a role for our industrial -- our industry partners develop some of the great technologies that we have, and then our academic research is also important. So, we just hassle those things out as a debate; is this a reputable and responsible use? That's a real awesome responsibility that we take very seriously, and it could be occasionally controversial within the committee, but we try to just make sure to be stewards for anyone that has a use that could advance public health.

DR. PERKINS: Thank you. And the second question is under Proposal 2, relative performance against a clinical endpoint comparator that includes histology, I'm wondering,

just because it seems that other molecular tests may at some point replace histology, whether you want to use that specific wording in this proposal, because then if histology gets replaced, you run into the same problem of this being outdated, whereas if you use language that might be a little bit more vague, it could allow for histology to morph into the other molecular tests.

DR. ZARITSKY: That's a great point that we haven't thought of, so thank you for bringing that up. That's an extremely valid point, thank you.

DR. VAN DER POL: Okay, before you walk away, there is one last urgent question.

Dr. Thomson.

DR. THOMSON: I don't know how urgent, but it's important to me. In the composite comparison, have you considered having an analytic control, an engineered common analytic control as a starter?

DR. ZARITSKY: We haven't considered that, but that is definitely something that is -- we should think about. I don't know. Okay, maybe -- yeah, because I don't know what kind of analytic controls are actually out there. Maybe you can --

DR. SCHIFFMAN: We tried clean prep clones and different preps of adding different cellular material and interferers and everything into those to try to get some kind of common reference reagent that could be in bulk and could be a standard. It's really hard because the composition of the glycans of the cervical mucus, the interference of heme from blood, the consistency of the sample, the inflammatory reactants make clean preps very easy to get right. So, it's a good first step. If you get the -- if you miss the clean preps, which are just composed DNA of the right types, you're in trouble. But I think I'm being told to stop.

(Laughter.)

DR. VAN DER POL: Thank you, everyone.

We're now going to take a break. We're going to keep that to 10 minutes, if possible, 15 if you really can't get back here, but try to get back as quickly as you can so we can catch up on time. Oh, and please don't -- the panelists, please don't discuss any of the things that we've heard here this morning amongst yourselves or with anyone else while you're on break.

(Off the record at 10:38 a.m.)

(On the record at 10:58 a.m.)

DR. VAN DER POL: Okay, we're going to proceed with the Open Public Hearing portion of the meeting, and the public attendees will be given the opportunity to address the Panel to present data, information, or views relevant to the meeting agenda. Ms. Asefa will read the Open Public Hearing disclosure process statement.

MS. ASEFA: Both the Food and Drug Administration and the public believe in a transparent process for information gathering and decision making. To ensure such transparency at the Open Public Hearing session of the Advisory Committee meeting, FDA believes that it is important to understand the context of an individual's presentation. For this reason, FDA encourages you, the Open Public Hearing speaker, at the beginning of your written or oral statement, to advise the Committee of any financial relationship that you may have with any company or group that may be affected by the topic. For example, the financial information may include a company's or a group's payment of your travel, lodging, or other expenses in connection with your attendance at the meeting. Likewise, FDA encourages you, at the beginning of your statement, to advise the Committee if you do not have any financial relationships. However, if you choose not to address this issue of financial relationships at the beginning of your statement, it will not preclude you from speaking. Finally, if any speaker is reading for someone else, please state this at the beginning of your statement as well.

FDA has received five requests to speak prior to the final date published in the *Federal Register*, and each speaker will be given around 10 minutes to speak.

DR. VAN DER POL: The first speaker that we have today is Dr. Jeffrey Andrews. Please come forward to the microphone. We ask that you speak clearly to allow the transcriptionist to provide an accurate transcription of the proceedings of this meeting.

DR. ANDREWS: Hello, my name is Dr. Jeff Andrews. I'm a OB/GYN and colposcopist. I'm the worldwide medical director for BD Diagnostic Systems, so I'm a full-time employee of BD Diagnostic Systems, BD Life Sciences, and BD Global Health.

Dr. Van Der Pol, thank you for recognizing me, and I appreciate the opportunity provided by the FDA to speak today. Today we celebrate International Women's Day, and everyone here is contributing in important ways to women's public health and secondary prevention of cervical cancer, working toward the WHO's goal of elimination.

Simple detection of HPV infection is not enough to protect patient safety and perform effectively, so HPV assays must accurately detect infection and establish performance relevant to a patient's risk for precancer; therefore, the assay has to be applicable to the screening population and cannot be biased either toward detection and infection at the expense of identifying precancer disease risks.

As we heard, HPV assays are becoming increasingly important due to adoption of HPV primary screening because of high assay sensitivity and lengthening of patient screening intervals that go along with the improved negative predictive value. Also, the increase in vaccination rates and shifts in genotype prevalence are important.

So, you've asked these questions about reducing the validation burden for HPV assays and maintaining patient safety protection. We think, at BD, that it's very important that the study population be the same as the intended screening population, so we have the rhetorical question about how we can assure the right mix of non-disease and disease

samples within the study population and how to assure safety for future screening populations with any new device. So, HPV assays need to perform well in the general screening population and not just in a particular subset population.

The study population needs to be representative of the screening population and evaluable by range of age groups, reflect the cytology differences across the screening population, have different target HPV viral loads near the clinical cutoff, as well as the screening history and other factors. And we know that the screening population itself is undergoing change.

Negative predictive value is the critical primary screening metric for evaluating assay performance with respect to safety, and that mandates requirements for either significant number of screen negative subjects undergoing colposcopy at baseline endpoints and/or 3-year longitudinal data and/or well-characterized biobanks of residual samples. The use of biobanks and well-characterized archived samples are reasonable options provided they adequately represent a screening population.

Limiting the proportion of vaccinated subjects to increase the prevalence of disease creates a conundrum. Capping must be accomplished in a manner that allows sufficient statistical power to understand performance in future highly vaccinated screening populations.

Also, samples cannot be solely derived from a referral population. In particular, viral loads associated with high-grade CIN are different than those in a screening population, and this is one example of that in a paper by Xi and also, as was stated by Dr. Zaritsky, you can see higher viral loads with higher-grade cytology and lower viral loads with NILM.

And we ask what considerations are crucial when contemplating a least burdensome clinical study design?

As was already pointed out, a simple molecular comparator is problematic, and

especially if you're considering infection versus the CIN3 endpoint. Not all comparator assay designs and outputs are the same due to different consensus primers and genotype-specific primers and probes, and the potential for establishing performance as compared solely against comparators that's acceptable, indicating detection of infection but non-clinically relevant lacking relation to precancer.

There's no apparent way to establish negative predictive value performance metrics to inform how well the assay will perform in a primary screening environment, and also no apparent way from this comparator method alone to inform clinicians about other key performance metrics: specificity, colposcopy rates, and longitudinal performance. I do note that Dr. Zaritsky mentioned there will be other studies needed to establish clinical cutoffs, so these may address some of these challenges.

The augmented molecular comparator with clinical endpoint is an improvement over the simple molecular comparator. Use of histopathologic information in conjunction with a molecular comparator would improve assessment of performance risk, and BD thinks this is acceptable. Inclusion of histologically defined disease precursors as a component to a molecular comparator improves the ability to assess clinical performance risk. And I would agree, I think it was Dr. Perkins that mentioned we may need to generalize this language to other methods of determining the state of disease other than histopathologic standards.

Even now, histopathologic reference standards have evolved, and those changes are critical to understanding what is a comparator positive and a comparator negative. Enhancements that have been used to date include the p16 biomarker and microdissection with PCR on the actual lesion.

As histopathologic science continues to evolve, HPV assays should be validated against the best clinical endpoints used by the medical community at the time of the study, and as was already suggested, it's even possible that histopathology could be replaced by

another new technology.

This is an article by Dr. Darragh, as part of the LAST criteria, which just demonstrates by Kaplan-Meier curve that those women with CIN2+ and -- sorry, CIN2 and p16 positive went on to have a greater risk of disease than the women who were initially CIN2 and were p16 negative. We used the p16 biomarker as the pathology in the Onclarity clinical trial, and we'll be publishing our results in the near future.

Multiple opportunistic screening paradigms and management approaches exist. We have liquid-based cytology with optional triage to HPV in the U.S., which discriminates risk between those who are positive and go on to colposcopy and those who are negative and return to screening.

We have primary HPV and triage of positives to improve the positive predictive value of just the HPV positive result, and we're currently mostly using liquid-based cytology and/or partial genotyping (16, 18) for triage. And then there's co-testing with sorting by both cytology and partial genotyping.

In the future, we may see extended HPV genotyping beyond Types 16 and 18 to discriminate risk. We may see immunohistochemical dual-staining cytology to discriminate risk, molecular biomarkers, and epigenetic marker panels such as methylation tests, and also, it's possible, even likely, that we'll see screen-triage-triage strategies, which is even more complicated.

Overall, BD supports consolidation of indications while recognizing, as I think Dr. Schiffman was saying, that primary HPV screening is the way forward, endorsed by WHO, ASCO, ASCCP, and that the standard for the consolidated indication should meet the minimum standards for primary HPV screening.

Risk-based guidelines by ASCCP and NCI are necessary, and we anticipate those. And critical patient management information includes what genotype the patient has and how

long it's been persisting, as was discussed by both Dr. Zaritsky and Dr. Schiffman.

Different options exist now for triage of HPV positive results, and more are on the horizon. HPV assays that report results for genotypes beyond 16, 18, and 45 align well with triage screening strategies that leverage the differential oncogenic risk of the HPV genotypes. The ability to utilize an extended genotype assay design in a diagnostic environment is dependent on several factors: clinical practice guideline developments as well as what's being discussed here.

Safety and effectiveness should remain the priority. Test samples should be representative of the intended use population. Assay outputs must be validated as being clinically relevant to the disease precursors, as they serve as inputs to risk-based screening strategies.

The effectiveness or positive predictive value of the HPV test result is improved by triage, but the screen-triage or screen-triage-triage may be uncoupled for regulatory purposes, and BD supports removing references to specific triage tests and clinical actions.

The impact of vaccination has already been discussed, so I've just reviewed some of the points there in the interest of time.

Diagnostic use of HPV assays that report extended genotyping results will allow the clinical community to utilize real-world assay outputs to evolve screening guidelines and improve the ability to triage patients and discriminate risk categorically.

And I thank you very much for your time.

DR. VAN DER POL: Thank you, Dr. Andrews.

Our next speaker is Ms. -- is Professor Cocuzza from Copan.

(Pause.)

DR. VAN DER POL: You might hit escape and then you might see it.

DR. COCUZZA: All right. Thank you.

Thank you very much, and I'd like to thank the FDA and the Committee for allowing me to present some of the preliminary data from our study we've been carrying out and validating and standardizing the work for HPV testing on self-collected samples. I actually work for the University of Milano-Bicocca, and I'd like to thank also Copan for sponsoring my transport to attend the meeting. I have no other conflicts of interest.

So, to start with -- okay, sorry. Okay, so it's well known that cervical cancer still plays an important disease burden in spite of improved measures in primary and secondary prevention. However, most cervical cancer cases that occur nowadays are actually associated with women who do not participate to screening programs and under-screened women, and this has called for an action from the international organization to improve strategies for cervical cancer prevention. So self-sampling offers a promising alternative in cervical cancer screening, so particularly as it has been shown to improve participation of hard-to-reach women to screening programs.

However, the accuracy of HPV testing on self-collected samples needs to be evaluated prior to their introduction in screening programs. And so, a first step forward was actually the meta-analysis performed by Mark Arbyn and colleagues which actually showed that only PCR-based HPV tests can offer a similar accuracy on testing self-collected samples.

More recently, another protocol has been published, the VALHUDES protocol, which is aiming at evaluating accuracy on testing self-samples, considering not only the validated HPV assay but also the collection device, so going to a standardization from sample collection to HPV testing.

So, as I said, introduction of self-sampling in screening programs requires validation of the procedure and, in particular, the standardization not only of the HPV assay but the collection devices as well as what comes in between, such as the nucleic acid extraction

procedure which many assays nowadays have included in HPV testing.

So, challenges are certainly the avoidance of false negative results, and as already shown by the meta-analysis, assays which have higher sensitivity, analytical sensitivity, offer better accuracy in testing a self-collected sample. But, also, validation of sample adequacy from self-collected sample with specific collection devices is necessary.

This introduction of validated procedures, standardized procedures, is made even more difficult by the fact that most manufacturers do not include self-sampling as intended use in the assay package insert, which makes it difficult for the laboratories and screening organization to use this protocol.

And this is evident in the table. I'm not sure whether it is clearly readable, but you can see different studies addressing vaginal self-collection in different countries, and you can see that there is a wide variety of collection devices that are being used, as well as a wide variety of media for resuspending the dry vaginal swabs, not only in terms of the media but also of the volume used as well as great differences in the starting volume for nucleic acid extraction and elution volumes as well as HPV assays. So, a lot of standardization is lacking, making difficulty in comparing results from different studies.

So, we wanted to perform a pilot study aimed at evaluating HPV detection on self-vaginal and urine samples collected with specific devices, as compared to clinician-collected sample, which is our gold standard, by means of standardized sample procedures. In particular, we focused on evaluating sample adequacy in terms of cellularity, concordance in HPV detection from between self-collected samples and clinician-collected samples by means of two different clinically validated HPV assays commercially available with different levels of automation. Moreover, we tried to standardize the resuspension volume in order for it to be in keeping with commercially available media in terms of both type and volumes in order to make this process high throughput.

So, this is the study protocol which I would like to show you some of the preliminary results. The study protocol is similar to that of the VALHUDES study recently published by Arbyn, in which women are enrolled at colposcopy. A hundred women were considered in this primary analysis, and they were asked to collect, at the time of colposcopy, vaginal samples using a dry vaginal swab, flocked swab, urine samples using two devices which we compared, which were Colli-Pee from Novosanis and UriSponge from Copan, and prior to colposcopy they also are cervical samples collected by the clinician.

We first tested the samples using a full genotyping HPV assay, Anyplex HPV28 from Seegene, which is not a combined pre-analytical phase, so we defined the pre-analytical phase starting from 1 mL of each sample type, extracted with NucliSENS easyMAG, and then eluting in different volumes, 100 µL for cervical and vaginal samples and 40 µL for urine samples.

But prior to showing you the results of testing, I'd like to show you the sample adequacy assessment which we performed on this hundred sample types which used a real-time quantitative PCR assay for detecting a double-copied human gene present in the cells. And we've seen the stable average total number of cells per sample, which you can see shows very comparable results between all the different types, taking into account that this is the total number of cells per sample. If we look at the range of the cellularity, as expected, we see a wider range for urine samples as compared to cervical and vaginal samples.

If we move on to the concordance in HPV detection between the vaginal/urine as compared to the cervical sample, we see that the overall positivity rates that we observed in these samples were very similar, particularly on the red part of the graph where you can see high-risk types, one or more being detected with similar positivity. And also, in terms of genotype detection, we observed similar genotypes, you can see here in orange, in the

cervical cells in the vaginal samples and in the urine. So, again, a good concordance was observed from the sample types, as shown in this chart.

If we look at the -- we are not able with the sample size to define clinical sensitivity and specificity, but we were able to compare some of the clinical results obtained from the women enrolled in the study. Involving the chart, you can see the cytology referral grade, and below those 22 women who underwent biopsy, the grade is shown as histology, and here you can see the cytology/histology as compared to HPV detection for all the three sample types analyzed. And you can see a very good concordance was observed.

And interestingly, below, the last two samples showed -- were referred for high SIL but were noted to have a negative biopsy, and reliably, the HPV testing was negative for all sample types demonstrating that HPV testing is more reliable than cytology in these two patients. So, agreement was observed; as I mentioned, it was very good for the high-risk types, particularly for 16 and 18, which were more numerous in this population studied.

We also compared urine sample collection devices in terms of HPV detection; again, very similar percentages were observed for the two different urine samples collected with the two different devices. And, also, this was confirmed by HPV genotype detection and concordance with cervical samples.

On a subset of samples from the first 60 women, the samples were aliquoted and stored, and they were used for further testing using more automated FDA-approved HPV testing, which was associated with a standardized protocol automated on the BD Viper system. Here, we're not able to define any of the aspects in terms of sample volume, which was defined by the company as being 0.5 mL, and the process of nucleic acid extraction and testing was automated.

There was, however, software analyzing cutoff values which were previously validated on cervical samples, but we decided to overwrite the cutoff values set by the

software because we were analyzing different sample types. And in doing that, we observed again a very good concordance in positivity rates between the three sample types using a validated system which has associated standardized workflow for the pre-analytical side.

However, if we interpreted the Ct value using the software, we noticed that we would have missed some HPV positivity, but -- vaginal sample emphasizing that cutoffs, clinical cutoff values may need changing when analyzing self-collected samples.

We also tried to standardize -- I'm nearly finished, one more last slide -- the volume in order for this to be a high-throughput system, and we found that even if the sample was eluted in 20 mL, the concordance for the HPV detection was still very good.

So, in conclusion, the results that we obtained from this preliminary study in terms of HPV detection, I must emphasize, were very good in terms of cellularity, but we need to have a bigger study to clinically validate specificity and sensitivity. So, in the second part of this year we will be performing a European VALHUDES study, which will involve three different countries collecting samples in this colposcopy setup as the study I've just presented.

Thank you for your attention.

DR. VAN DER POL: Thank you.

We'll now hear from Keivan Eitefagh from CellSolutions.

DR. EITEFAGH: First off, I would like to apologize for having a cough drop in my mouth. As we celebrate women today, I have a wife that is a pediatrician and three children, three girls under the age of 6, all in either kindergarten or daycare, so therefore it's a necessary precaution in order for me to get through this speech.

Like I was introduced, my name is Keivan Eitefagh. I'm with a company, a small company in Greensboro, North Carolina, called CellSolutions. I am not a cytologist or a

molecular biologist or virologist or an M.D. In fact, I'm probably the most ignorant person in the room, which allows me to learn a lot today. I do have my doctorate in toxicology with a focus on analytical chemistry, so I know a little bit about statistics and detecting small molecules and that sort of thing. And even though my wife says that's being replaced with Disney princess knowledge, I still have my notes here in case I need a reference. So, before the speech goes into a tale as long as time, let's get down to business, and thank you for letting me be part of your world.

So, I'm going to focus a little bit on something that -- we're talking a lot about medical devices, and something that often gets ignored or not talked about, I think, are the preservatives that go into medical device approval. So, it requires the same burden of approval that the FDA requires for a new detection of HPV, a preservative. It's very cost prohibitive right now. Like, it's been echoed some of the difficulties listed before. It costs a lot of money, millions of dollars, you need tens of thousands of samples, it's very time consuming, and finding the willing clinical trial sites can be a big burden, especially for small companies like mine.

And one of the biggest problems we face in at least getting preservatives approved is the requirement of alternate collections. So, you either have an alternate two-sample collection from the patient in which you are comparing not really apples to apples, those two sample aren't equivalent and therefore we need to take a lot more of them in order to represent the population and be statistically relevant.

So, our goal is to sort of propose that we accept approved tests that already are FDA approved as a standard for comparison, and if we can establish a same equivalent sample as opposed to a heterogeneous Sample 1/Sample 2, then we can standardize the number of cells collected because we're collecting only one sample as opposed to two. We can standardize the number of theoretical infected cells or the degree of the infection or

disease and the viral load of those cells. So, therefore, when we collect only one sample instead of alternating samples, we can perform a more robust statistical analysis using a kappa statistical analysis or a non-inferiority case.

So, the reason why I'm here and not someone more applicable to this field is that most of my peers that are in the fields of cytology and molecular biology are actually performing a study over in Europe in which they are collecting this data.

Traditionally, we need heterogeneous comparison using two sample collections. This is an apples-to-oranges scenario, and it's a prolonged longitudinal study with required biopsy for comparison. And what I'm proposing is dividing one sample into two, one homogeneous sample into two; that way we can compare apples to apples and making them statistically relevant. This eliminates a lot of the cost and the number of samples needed. And if we can compare testing of already existing FDA-approved testing, such as target amplification, then dividing that homogeneous sample really doesn't impact the overall results.

So, this has already been proven successful. Over in Europe we have a group of pathologists running a myriad of samples that they collect. Luckily, fortunately, there's a high propensity of positive HPV population. We had actually, when comparing preservatives, a perfect agreement of 90 samples where both sets of preservatives, the predicate approved FDA preservatives and our preservative, had perfect agreement in detecting the positive samples in that study. And when we're comparing two already established tests that were both DNA-targeted amplification, I won't name names, I'm trying to -- but using that same analysis, using that single sample divided into two, you still had strong agreement with a 42% positive population.

And that's all I have to say.

DR. VAN DER POL: Thank you.

We'll next hear from -- we'll hear from Ms. Srinivasan -- apologies if I've messed up your name -- from the National Center for Health Research here in Washington.

DR. SRINIVASAN: Good morning. Thank you for the opportunity to speak today. My name is Dr. Varuna Srinivasan, and I am a physician with a master's of public health from Johns Hopkins University. I'm a senior fellow at the National Center for Health Research, which analyzes scientific and medical data to provide objective health information to patients, health professionals, and policymakers. We do not accept funding from drug and medical device companies, so I have no conflicts of interest.

We would like to start out by saying that we agree with the FDA's evaluation that a clinical endpoint of CIN3 for an HR HPV test is more meaningful in assessing precancerous lesions. We agree that although the prevalence of these lesions is seen in fewer numbers in the population, recruiting more women for the evaluation studies to detect CIN3 will help maximize resources and prevent overtreatment in the future.

For these reasons, the National Center for Health Research encourages the Panel to consider the presence of CIN3 alone as a positive result and not the combination of CIN2 and CIN3.

We disagree with the proposed indications for use. As the FDA pointed out, 90% of all HPV clears on its own. On average, women over 30 tend to have fewer sexual partners and more monogamous relationships than women under 30, and women over 65 had even fewer. That is why the USPSTF recommends co-testing with Pap smears every 5 years for women under 30 years of age. The FDA proposal to screen all women from 25 to 60 years of age with the HR HPV test exposes a large population of women to this test unnecessarily. For that reason, we believe that the HR HPV testing should be recommended only for women over the age of 30.

It is important to keep in mind that Pap smears directly determine the presence of

cancerous or precancerous cells in your body. Co-testing can provide the added benefit of identifying high-risk HPV infections and allow for more vigilant follow-up if HPV 16 or HPV 18 is diagnosed. We respectfully urge the Panel today to consider these suggestions while providing their final recommendations.

Thank you, and happy Women's Day.

DR. VAN DER POL: Thank you so much.

And our final speaker will be Dr. Evantash, representing Hologic.

DR. EVANTASH: Great. Thank you, FDA and distinguished members of the Panel and colleagues in attendance. I am Dr. Edward Evantash. I am an OB/GYN, and when I previously practiced, I was a colposcopist and a Chief of General Obstetrics and Gynecology at Tufts Medical Center in Boston. I'm currently the Vice President of Global Medical Affairs at Hologic, as my conflict of interest, as a full-time employee. And as many of you know, Hologic has been the industry leader in the goal of pursuing eradication of cervical cancer through innovation and diagnostic tools for early detection of disease.

We are especially interested and pleased by the agenda chosen for discussion today by the FDA, as it recognizes the shifting landscapes in cervical cancer diagnosis, as it relates to an increasingly vaccinated community and a greater of understanding of the role of HPV.

With these concepts in mind, there is a recognized need to harmonize the FDA labeling for diagnostic tools with the frequent updates and changes to the professional society guidelines, promote innovation to improve diagnostic technologies, and avoid unnecessary burden that is for government and industry, to achieve these goals without compromising the patient care. These are all worthy aspirations that Hologic strongly supports.

It is with this in mind that we have concerns with the basis for Question 4A, which seeks to establish safety and efficacy for new high-risk HPV devices by agreement against a

composite molecular comparator of three FDA-approved molecular assays. We agree with the FDA's observations that are outlined in the Executive Summary. Sensitive detection of HPV nucleic acid, both mRNA and DNA, has been established by the existing assays on market, and HPV diagnostic assays designed to be more specific may have discordant results with a composite comparator. These results may be incorrectly labeled as false negative when, in fact, the underlying infections might likely be transient and not associated with CIN3+ disease, or even non-oncogenic HPV infections.

Innovation focused on maintaining acceptable sensitivity and improving specificity to avoid unnecessary interventions associated with overtreatment would essentially be stifled in order to improve performance through greater detection compared to the approved assays. Without histologic assessment for disease, it would be difficult to assess the clinical relevance of any discrepancies between the investigational device and the comparator; therefore, we strongly support the proposal described in Question 4B and are interested to learn more about the opportunity to use a mixed clinical endpoint as described in Question 5.

Again, thank you to the FDA and to the panelists and participants. I yield the remaining time back to the Chair. Thank you.

DR. VAN DER POL: Thank you, Dr. Evantash.

This now allows us an opportunity before lunch to open up the floor to any people in the audience that would like to ask questions or make a statement.

(No response.)

DR. VAN DER POL: If there are no questions, then we've gotten to lunch early. You gave us a little bit of a tease there.

(Laughter.)

DR. VAN DER POL: Okay, I now pronounce the Open Public Hearing to be officially

closed. Thank you for your attendance.

For those of you on the Panel, there is a restaurant that will have a buffet lunch, which will get you in and out. We will reconvene starting -- we're going to start at 1:00? Yeah, so let's try to come back at 12:45, which gives you an hour for lunch and gets us moving a little bit earlier. Do take your belongings with you because you won't be allowed back into the room once they lock it up, so take what you need with you. Also, for those of you on the Panel, if you have not already arranged with -- is it Maryann?

MS. ASEFA: AnnMarie.

DR. VAN DER POL: AnnMarie Williams at the desk at the table outside about your transportation to the airports, please stop by and do that now so she has an accurate head count of who needs transportation. Okay, we'll see you in an hour.

(Whereupon, at 11:38 a.m., a lunch recess was taken.)

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

(12:48 p.m.)

DR. VAN DER POL: Welcome back, everyone. We're going to now call this Panel back to order, and we're going to focus on the panel deliberations at this point. Although this portion of the proceedings are open to public observers, public attendees may not participate except at the specific request of the Panel chair. Additionally, we request that all persons who are asked to speak identify themselves each time; this helps the transcriptionist correctly identify the speaker.

Before we proceed to the panel deliberations, I would like to ask our non-voting members, Ms. Joy Hugick, our Consumer Representative; Mr. Bradford M. Spring, our Industry Representative, if they have any additional comments that they'd like to make.

MR. SPRING: Just that as we go through the deliberations and discussions today and focus on the various options that FDA has put forward, just keep in mind that the scope of the discussion would be focused not just on new devices, right, that we also have to consider modifications to existing devices, those modifications being a new indication for use, new populations, new specimen type. So when we look at, for example, the various options around study designs, just keep in mind this is for a new device or a modified device because the answer may be different on the path we choose.

MS. HUGICK: Thank you. This is Joy McVey Hugick, the Consumer Representative filling in today. I would say, especially because the Patient Representative isn't here either, that I'd just ask the Panel and FDA to be cognizant of the patients who are participating in the clinical studies, especially when it comes to preventing overtreatment that might not be necessary, and at the same time making sure that, in making these decisions, that we don't want to compromise patient care. I do appreciate the suggestions for keeping more vague and broad language to handle the changing landscape over time and was glad to be hearing

that comment this morning by Dr. Perkins.

DR. VAN DER POL: Thank you, both.

Is the FDA prepared to answer any questions that the Panel may have on the topics posed this morning?

DR. FELDBLYUM: Yes, we are.

DR. VAN DER POL: Thank you.

Okay, now Mary Barcus will read the questions back to us to remind us of what we're actually talking about. I do want to say that we're going to change the order up a little bit because Question 4A, in particular, depending on what people think about that and what the consensus, if any, is about, that could influence conversations around Questions 1, 2, and 3. So it seems like, perhaps, if we start with 4, we might be a little bit more organized going forward.

DR. BURK: In the document it mentioned if there were other suggestions or proposals, and Dr. Thomson had begun to mention the issue of an analytical test. I don't know at what point we might want to discuss that, having an analytical HPV test that would be merged in with possibly some of the other three.

DR. VAN DER POL: Right. So as we go through each question, that will be your opportunity to make suggestions just like that and then to have the Panel discuss that and also then get feedback from the FDA about their thoughts on that.

DR. BARCUS: Okay, so the first question that we're asking for input on is: Would the Panel recommend one or a combination of the following three proposals to increase the number of women positive for CIN3+ and/or HR HPV in clinical studies:

1. Supplementing from referral clinics
2. Utilizing archived specimens
3. Capping the vaccinated population

The second question for the Panel: Regarding the NILM/HR HPV double negative and ASC-US/HR HPV double negative populations in clinical studies supporting HPV device approval:

Do the benefits of colposcopy referral for the assessment of verification bias outweigh the risks associated with the procedure and potential overtreatment?

And we'd ask that you please discuss for each of the two populations separately.

Question 3: Regarding the indications for use (IFU), do the benefits outweigh the risks for:

A. Consolidating the indications to encompass one general screening population
B. Removing references to specific triage tests and clinical actions

And, again, please discuss any potential risk mitigation measures if a new IFU statement were to be used.

Question 4: Please discuss whether the following types of data evaluations are acceptable for the assessment of safety and effectiveness for new HPV devices:

A. Adoption of a molecular composite comparator method
B. Evaluation of relative performance against a clinical endpoint comparator

And please discuss minimum acceptable performance criteria.

And the fifth question is: If the Panel recommends assessing HPV device performance against a clinical endpoint comparator:

A. Is utilizing a mixed histological/molecular comparator acceptable?
B. If so, how should the combination of HPV result and histological diagnosis factor in when assigning "comparator positive" and "comparator negative" results?

And I'll just go back to Question 4 for the initial discussion.

DR. VAN DER POL: Thank you.

So our first topic will be, does the Panel -- the Panel should discuss the following types of data evaluation that are acceptable, and so this really relates to and will drive study design, obviously, because this is going to define the endpoint. So we need to discuss Option A, which is the adoption of molecular composite comparator method, and separately, as an alternative, Option B, which is relative performance against a clinical endpoint comparator. I would suggest there may be an Option C, which is to say that you're going to use molecular comparators with endpoints with follow-up visits based only for those people who have discordant results. So I think there is not necessarily A versus B, but there may be some continuum here, and so I'd like to hear from the Panel on their thoughts about all of this topic.

DR. MOSCICKI: Barbara Moscicki, UCLA.

I had a question very specifically in regards to I have a hard time with the molecular composite because that's going to be difficult for really young age groups, especially 21 to 24, where that's currently not even a practice and that we don't know what even 25 to 29 would be in a vaccinated group.

The other piece that I was struggling, I don't know where this goes into, but really talking about the prospective studies that are going to be needed, and we haven't discussed that because, really, what we've all calculated and was presented earlier today was what is my 3-year risk, what is my 5-year risk after a negative test, and I think if we don't have some type of sense of what that negative test does over a long period of time, I think it would be difficult to assess even intervals for a new test.

DR. WENTZENSEN: I think --

DR. VAN DER POL: Don't forget your name when you're starting. It's hard for the transcription guys.

DR. WENTZENSEN: Nicolas Wentzensen from NCI.

Regarding the question of a molecular composite comparator method, it would be extremely important if that was done to make sure that the sample set -- or the sample set really reflects the whole spectrum of disease, and I think that is -- that's a crucial point. If you only take cases you will have a very hard high viral load, a situation you -- it's much easier for assays to agree, whereas if you include more HPV positive NILMs, you may have a lot more discrepancy. I think putting this comparison of that together requires histology in order to really understand what you're comparing, and if you have that, you're almost automatically in a 4B situation. I think you cannot do this without having histology information.

DR. DARRAGH: Teresa Darragh, UCSF.

I actually have maybe a potential for a different paradigm, that when you're looking at a composite molecular comparator, you're really looking at the true negatives, and we all need to accept the exquisite negative predictive value of a negative HPV test. So using, in a study design, using a composite molecular comparator for the negatives, I think, makes a lot of sense.

And then for the women, the true positives, the women who actually have disease, then you could use, you know, your enriched sample, you know, the colpo clinic and the lead population and maybe archived specimens, and so you're having two different arms of your research design. One is looking at your true negatives, and one is looking at your true positives.

DR. VAN DER POL: Can I ask a question because I think you guys are saying different things, and so what I've heard is that some of the Panel members feel like every -- or every assay needs to develop its own prediction models based on negative predictive value over a period of time, while others are saying that a negative is a negative, if I can simplify what you're saying. And so I think that that's worthy of discussion because, to me, that is exactly

the crux of the matter, is where do we fall in that kind of question.

DR. BURK: Robbie Burk, Einstein.

Well, I'd still like to bring up the issue about the quantitative detection of different HPV types and potential variants in a given sample. At the end of the day, that's really the critical matter. We've moved forward in our knowledge about the natural history of HPV and cervical disease. So, you know, there are issues of detecting HPV in different sample types, which probably should be validated within the given material, but at the end of the day, as we move forward, it's really about detecting and quantitating HPV in the given sample and what that risk is from our knowledge of natural history. So I would -- you know, maybe with A, so molecular composite and analytical sensitivity should be considered.

DR. FELDMAN: So I think maybe we're thinking that it shouldn't be a dichotomous decision about what type of study you use as well as what type of comparator you use for every single situation. So building on what Terri was saying, from my perspective, you can start with a certain sort of sample. Let's just suppose you started with a very enriched sample and it failed to show you a positivity where it should, right?

Right then, from my perspective, if you didn't show a positive test when you should show a positive test in a highly -- like a colpo clinic where you're going to have pathology, so you're going to have a histologic endpoint, and you know the viral load should be pretty high, and you know that if those women have it, it should be detectable. So if the test doesn't show positivity corresponding to the -- you know, the CIN diagnosis as well as the other HPV tests, in that situation, if it fails there, from my perspective it doesn't then go on to be validated in a lower-risk population. When you get to a lower-risk population, let's say you have samples of Paps from a routine screening population, so it's sort of a much lower-risk population, you might, at that point, once it's passed through the first one, be able to more accurately and comfortably go with determining if it can pick up abnormalities

in that using only a -- like a marker test, not needing the clinical validation.

So, to me, I think you can use the studies for different aspects of the question, and I think, on to your point, we need to know both the true negatives, and we also need to know the true positives, and those might require slightly different studies.

DR. PERKINS: This is Rebecca Perkins from Boston University.

I have concerns about A, the adoption of just the molecular composite, because I think that would assume that the current tests are perfect and it would potentially -- and cannot be improved upon, and it would potentially disincentivize advances related to clinical endpoints, so I think that's important to consider.

DR. MASSAD: Leslie Massad.

So the other issue I have with A is that if you're validating against previously validated tests rather than against a clinical population, you're looking backward at an essentially pre-vaccination population, and you don't show validity against contemporary screening groups.

DR. KINNEY: Walter Kinney.

Is it possible that screening intervals are an issue of clinical practice that the FDA doesn't necessarily want to get involved in? And that they may change at the whim of professional societies?

DR. VAN DER POL: I think that's perhaps one of the questions that will come up in one of the others when we talk about intended use. So right now we're just really kind of focusing on what the standard should be for comparison, and should it be a molecular composite, or does there have to be a more clinical endpoint in order to really validate these assays as they're developed?

DR. UNGER: Beth Unger from CDC.

I have to agree that it would be ideal to move to an analytic endpoint and include

some analytic parameters within the test, but I think we don't fully understand how the clinical samples impact the efficiency of each assay. So it is true that a plasmid-based, you know, quantitative assay is unambiguous as to where the analytic cutoff is, but in a real clinical sample, you have the problem of the sample preservation, the efficiency of extraction, and a lot of other variables. But having said that, I still feel like if we -- we're going to be faced with having to move away from clinical lesions, and so I would just suggest that we use a clinical comparator but also be sure that we include an analytic component to it so we have some comparison of what is the analytic threshold where these assays work.

DR. WENTZENSEN: Nicolas Wentzensen.

Beth, can you extend on that maybe and explain what the analytic endpoint would be?

DR. UNGER: Well, I guess I would be asking for twofold. One is establish the analytic performance for each of the assays using plasmid standards or some very defined component, and that, of course, would require that the plasmid be compatible with the assay format, which is one of the problems, and it also would require some other manipulation for an RNA-based assay, so that is a limitation.

And, in addition, I think that the material that's used in the -- in each of the assays, whether it's accrued, extract, or just the cells or whatever, be quantitated so you know quantitatively how much human cell DNA you're putting in as well as how much HPV DNA you're putting in and those assays can be done.

MR. SPRING: Brad Spring with BD.

Based on the comments and questions around analytical studies, I don't know if FDA can comment on what other studies manufacturers typically do beyond and in addition to the clinical studies because that may help answer some of the questions. There's a lot of

other analytical work that's done, and I don't know if, Tamara or Uwe, you can comment on that.

DR. FELDBLYUM: Tamara Feldblyum.

So every assay, including HPV and all other infectious diseases we look at, requires analytical studies and clinical studies, and analytical studies usually include limit of detection, they include -- okay? I got to find this. I wasn't heard by the transcriber.

They include usually limit of detection. They include cross-reactivity with other organisms that you potentially can find in anatomical sites where the sample is coming from. We also look at cross-reactivity if it is a molecular assay. Like, in this case, you look at primers, and probes don't cross-react with organisms that you're not supposed to detect. You look at exclusivity/inclusivity. You also look at instruments that the assay is performed on. For example, do you carry cross-contamination when you perform -- so there's a whole slew of analytical assays.

What we are discussing here is really clinical assessment; analytical will be done no matter what we decide here. So in this case of HPV, we have done and we have requested manufacturers to do some comparisons assay to assay similar to what is proposed here, but it's usually the new HPV assay versus one assay that's already approved, and that testing is done on a small subset of clinical samples, and that is considered analytical comparison. What we're thinking is using these molecular assays that have been approved more as a clinical comparator so the testing would be performed on all women recruited for a clinical study as one of the options.

DR. UNGER: But do you know the analytical performance in the clinical material?

DR. ZARITSKY: Yeah, so for the analytical comparator, so what we usually do now is, as Tamara said, take a subset of the specimens from the clinical study and then compare the investigational device result against what's called an analytical comparator, which

usually consists of either two FDA-approved tests or one FDA-approved test and PCR sequencing or some sort of molecular test, and then we assess agreement between the new test and then the comparator result. But we don't -- I don't know if you're asking more about like analytical limit of detection or something or -- and we don't do that.

DR. FELDBLYUM: But we do limit of detection, and it's done in different materials, you know --

DR. ZARITSKY: Yeah, but that -- so the limit of detection, it's -- I guess it's typically a C-95 because it's used -- the clinical cutoff is used, and we do look at 95% positivity when the clinical cutoff is used, not the analytical.

DR. UNGER: I'm just trying to make a little bridge between what we know about performance of a purer sample versus a biologic sample. So if we can understand how much the biologic components in the real samples influence the ability of each assay to detect, then we're moving closer and closer to being able to use an absolute standard.

DR. VAN DER POL: Right, but --

DR. UNGER: And so I'm just urging that we try to include that as a component, because in the end, it is the nucleic acids that we're looking at, and that can be pretty unambiguous.

DR. VAN DER POL: And I think that -- this is Barbara Van Der Pol.

And I do think that that's part of the analytical assessments that are done before actually going into the field by most manufacturers, and so I think we actually have that, and to me, the question here is, is it important? How important is it? I mean, you can phrase that any way you want, but is it necessary for manufacturers to do longitudinal studies because if you're going -- if you're really going to look at CIN3 as an outcome, you are not necessarily going to get that on Day 1's test, and so if you say that you want CIN3, you either have to enrich the population, which we can come back to, and it's one of the

earlier questions, or you have to say we're going to do head-to-head comparisons or, as I suggested earlier, which, you know, I'd encourage you to talk about and see if it makes sense as an option, is to say that all the agreed negatives are negative and we have a sense of that negative predictive value. Do we need to do further follow-up work with these people? Do we need to do colposcopy with these people?

All of the positives that were positive by the other tests, you know, we can find out if they're really positive, if they're going to result in colposcopy anyway at that time, but then do we need to do something like call back some people that have discordant results to try and understand why those results are discordant, because it may be that a new test is a better test and that's why you can't rely solely on a molecular comparator. But finding some sort of hybrid of those solutions seems to be like an option.

DR. WENTZENSEN: Nicolas Wentzensen.

Having the longitudinal, the prospective data, is important for the negatives to see how long we get reassurance from a negative test. I mean, that is kind of -- that is how all the primary screening studies have been conducted, and depending on how we select this composite comparator, if we are missing a group of women with an early HPV infection that is just not in a colpo population, then we may not be able to have that comparison to really address the reassurance, and I think that's why we -- like if this composite comparator would need to be set up like a full population that is undergoing screening, so you need to really characterize it well. You can't just take some mix of samples that is maybe the two ends of the spectrum and you're missing the important middle parts, so I think -- that's why I think knowing histology endpoints is important for this, and then once you have that, you can do the comparison to the clinical endpoints.

DR. VAN DER POL: Go ahead.

DR. MASSAD: So I would answer a question that you had just posed as, yes, we do

need further exploration of discordants when there's discordants between the investigative device and the established device but also when there's discordants between multiple established devices. In the paper that was circulated before the meeting, it says that when there's discrepancy between the established devices, those are excluded from analysis, where I would find that to be the population that's of really greatest interest.

DR. MOSCICKI: I'm being just also a little confused because it seems like we're overlapping into two. What do we do with the NILM HPV negative? So are we answering both questions here at the same time? Just a little clarification.

DR. VAN DER POL: No, I mean, I guess for this particular question, really, the topic here is does the Panel feel like it makes sense to have a molecular composite comparator method? What I'm hearing from most people is that that is insufficient.

DR. GRAVITT: Yeah, the way that I look at it is that the -- a lot of the established science is already there in terms of what we consider an HPV test to be with regard to its negative predictive value, and if you did as I think what Nicolas was suggesting, rather than saying you're going to have a micro-comparator just in a colposcopy referral population, what you essentially want to do is to have two different populations where you're going to do an HPV-by-HPV -- an investigational HPV test by a comparator in a general screening population, and you want to power that to say that you've got non-inferiority with just detectability because in that population you're going to have about 10% discordants probably between any HPV test or repeating the test twice with the same HPV, and then you want to see non-inferiority in the colposcopy referral population and the CIN confirmed histologic outcome population. And in some ways that also allows you, if someone did want to confer with a claim to say that they had equivalent sensitivity but their test had a superior specificity, which is something we all want, that allows you to use a similar study design, it would seem to me to show equivalence within the CIN2/3 population but then

showing significant discordants really within the normal population. So if you are equivalent in the CIN2+ but discordant in the direction that you're claiming in the general screening population, then that should be a pretty good indicator because that's essentially what the sensitivity and specificities and long-term follow-up are showing. And so I believe that, you know, it's a question of whether we believe that that is kind of known evidence, and so as long as we're not -- because it comes down to are you picking up the same HPVs?

DR. ALEXANDER: Barbara Alexander at Duke.

I think that there is another issue while we're talking about the specificity and the long-term negative predictive value. There's also that analytical sensitivity, and likely, in my mind, the way the molecular field is going, we get increased analytic sensitivity. And so with that increased sensitivity, without some kind of clinical endpoint, we're not going to know if we're negatively affecting by overtreating or appropriately treating, right, so default's positive, so to speak.

DR. VAN DER POL: Right, but if you did a study design that really had an in-depth discordant analysis and perhaps with follow-up visits for those specific patients, not the whole cohort but any patients that had initial discordant results between the comparator and the investigational device, then you would start to look into that, and then that would have to be how they set their cutoff as well, right? So it sounds to me, and I'm going to ask if you guys agree that I'm capturing the mood here, it sounds to me like 4A, in and of itself, is insufficient and that most people think that either 4B or some variation on 4B, and exactly how that looks, we're not quite clear yet, but I'm seeing a lot of nodding of heads.

DR. WENTZENSEN: I would like to add one point maybe. There are situations -- and so we're mostly talking about new assays coming along, but there are also situations where assays are slightly updated or whether they transfer to a different machine or things like that where previously there have been very big efforts in proving that they're equivalent.

This is a situation where I could see a -- potentially work, if it's a well-selected set of samples, like showing like minor changes that is, I think, really good for 4A but for a new test, I think we would like a little more.

DR. THOMSON: I would agree.

DR. BURK: In this comparator, how do we be sure that different populations that might harbor different HPVs associated with CIN3 will be adequately sampled?

DR. VAN DER POL: Well, but remember, this is not about how the study population is chosen. This is just how you're defining your endpoint for that evaluation, and so the populations would be chosen according to routine standard guidance, and I do think we come back to that in some of the other questions. But this is just once you've identified your population, how are you making that endpoint measurement, and it sounds like there are some cases where molecular comparators may be sufficient for a part of the population, and other parts of the population might need further examination and investigation.

One of the options as well, which I think -- if everybody's in agreement on this, I'm going to sum this up so we can go into 5 because that's going to talk about maybe sequencing being an option, which I think that's kind of the next logical step from here.

Okay, so just kind of to sum up our conversation about this, people are concerned that if the new test, the investigational test, is not compared to some sort of histological standard, that we won't necessarily be able to identify improvements in that new test. We also might not be able to identify shifting results based on a post-vaccination era, which I think was a valid concern as well, because I'm not sure that that's completely true because all of the new assays would be run on the new specimen, so I think that it would be probably a fairly even playing field, but even so, it was raised. And I think that people did have concerns that are worth noting about the age groups that would be included in these newer test evaluations and not maybe knowing how old or existing tests actually worked in

those age groups. So I think there were concerns about 4A, but I think 4B was found to be acceptable in some form or version.

Yeah, okay. Are there any other things people want to get into that before we go -- all right, let's go on to Question 5, then. So Question 5 says if the Panel recommends assessing HR HPV device performance against a clinical endpoint comparator, then there are two pieces here, and that is, is using a mixed histological/molecular comparator acceptable -- which is exactly what we were all just talking about -- and if so, how should the association of types with CIN2 and 3 lesions factor in when assigning comparator positive and comparator negative results? And this is the table that Luna had talked about earlier in the day that they had started populating and to give you a frame of reference for the conversation.

DR. DARRAGH: Teresa Darragh, UCSF.

So for my understanding, I think when we're talking about a histologic endpoint, there needs to be a couple of baseline things. So histology is used as the gold standard, but the gold standard isn't always quite so gold as you would like it to be. So I think it needs to be adjudicated pathology with the use of p16 or other biomarkers that help support a transforming HPV infection, and I think, for the molecular comparator here, you're not dealing with the device that we use in a screening situation, but you really need tissue-based microdissection and typing of the specific lesions.

DR. MASSAD: Just for clarity, that's for CIN2, not CIN1 or CIN3, or you would do p16 testing on all of those and microdissection for all of those?

DR. DARRAGH: I would do microdissection and typing on all the lesions because our type prevalence is rapidly changing with the introduction of vaccine and herd immunity, and if you're looking at a CIN3 lesion caused by either a virus not in the 9-valent vaccine or one that's not in the device makeup, then you would get potentially a false negative on that, so

you need to be comparing the apples and apples analogy.

DR. MASSAD: So in the figure you would change some of the positives based on the viral type?

DR. DARRAGH: Stu, I don't quite understand that question.

DR. MASSAD: So some of the 35, 39, 51, 56, 59, 68 CIN3s you might determine are actually negative?

DR. DARRAGH: No, the histology gold standard would say that it's a CIN3 caused by Type 35.

DR. PERKINS: Just to clarify, I think, Terri, you're saying that those people would show up in the cell that's CIN3 but testing HPV negative by the --

DR. DARRAGH: No, no.

DR. MASSAD: Or low risk?

DR. DARRAGH: No.

DR. PERKINS: But they would actually have HPV when you microdissected them, but in the assay they would show up as negative at this --

DR. DARRAGH: Potentially, yeah. Then the assay would not be sufficient to be able to detect those lesions. But you have to know what lesion --

DR. FELDMAN: Why it's missing.

DR. DARRAGH: Why it is missing them.

(Off microphone comment.)

DR. VAN DER POL: You're going to have to speak into the microphone so they can get the transcription.

DR. FELDMAN: This is Sarah Feldman.

So I think what you're trying to say, in order to make the gold standard more gold so that you really know what you're comparing, you would want both the p16 staining to make

sure, on the CIN2s, if you add -- they should be -- if they should --

(Off microphone comment.)

DR. FELDMAN: Or if they should be CIN -- right, in the selected endpoint of high grade or whatever versus a CIN1 if they were negative, and you also would want to know what HPV type was actually present in that lesion so that the test missed it and it was a CIN3, for example, you would say that test missed a CIN3, but we understand that it was because it didn't check for that HPV type. The one they test.

DR. DARRAGH: And just to continue, you have to know if something is called CIN3 by histology but doesn't have HPV by, you know, a PCR -- very sensitive test, that the gold standard may be incorrect.

DR. MOSCICKI: I think this is, again, one more argument for at least having some prospective follow-up on this group because we also know in colposcopy we miss lesions all the time, and therefore looking at other than saying it really was a false negative or positive, you need to look at what that lesion happened in 3 years, which is what we do now, we just say we missed the lesion. It's not a spontaneous lesion that happened 3 years later.

DR. VAN DER POL: So if you're trying to get an inferiority -- if you're trying to get an inferiority -- non-inferiority claim and all the other tests missed that as well, or got it, so you agree with all the other molecular tests and you disagree with histology, are you then non-inferior based on your ratios? And so I guess the question I keep trying to put to the Panel is we need to think about the burden on the patients as well, if we're asking these patients to stay into a study 2 or 3 years for what incremental advance in the diagnostic, you know, up front and I guess that's my question. It's not to say these patients shouldn't be followed. It's to say that how much more do we get out of understanding the diagnostic capacity of a new test by following for 2 to 3 years.

DR. MOSCICKI: I just wanted to clarify the question here because I don't see anywhere in here that says I'm looking at another FDA-approved test. We're just asking about mixed histology and a molecular comparator, so maybe I misunderstood this diagram.

DR. VAN DER POL: Yeah, the molecular comparator is also including, like if you're using approved tests, that's part of the molecular composite, and that was what the combination in B was to have the histology plus the comparator.

DR. FELDMAN: So I just want to comment on the burden on the patient. This is Sarah Feldman.

So if you say you're only going to do colpo at one point, then there's a likelihood the patient will get many more biopsies at that one point in order for the person who's trying to test it to prove that the patient has something, right? Because we know each incremental biopsy increases the sensitivity of colpo. But I can tell you, as a person who does a lot of colpo, that doing four biopsies is significantly more burdensome to a patient than a couple, than two or less. And two is actually sufficient most of the time, three or four adds a little bit more, but if you were the companies and you only got one time to do the colpo, you'd ask for four biopsies. It's way worse for a patient to undergo four biopsies from comfort than to say come back in a year or come back in 2 years or come back in 3 years and we'll, you know, recheck you and maybe do a biopsy at that point.

So from a patient's burden, I don't think time frame is as much of a burden as over -- the immediate moment where you overtreat them for teeny lesions that don't need to be treated, that wouldn't have been picked up with one or two biopsies, that sort of thing.

DR. VAN DER POL: But isn't that a treatment decision, not a study design decision?

DR. FELDMAN: The patient undergoing the study has to have those biopsies only at one point, and the biopsies themselves are very, very painful, so you don't want to do extra biopsies. So, you know, for example, one of the things about the cobas study or the Aptima

study is that it looks over 3 years, and so over time you're able to -- without doing excessive biopsies at the very first day, you're able to see over a certain period of time how many tests you need, how many biopsies do you need, how many lesions, you know, actually, you know, could've just waited, could've gone away, you might have missed them the first time, whatever, but there is real value to the patient of not being overburdened at that very first time with too much extra biopsy. That's really an unpleasant alternative.

MR. SPRING: This is Brad with BD.

We have to keep in mind, too, a lot of these patients have to consent to this as well, which also further reduces the population. So if they are getting anything extra, they have to consent to that, and we see this significant dropout of patients over time. So, you know, I'm hoping we're very clear up front to the patient on what this -- the risks are and what the potential burden is on the patient, too, but it's just something to keep in mind that the patient, I would hope, is consenting to this extra work, but that -- maybe your point is, Dr. Feldman, maybe they don't fully understand the true burden of having multiple biopsies taken in this case.

DR. BURK: Robbie Burk, Einstein.

So I just want to clarify that we're talking about CIN3 as the outcome because there's been some mention of CIN2, CIN3, and we heard from Dr. Schiffman the issues with CIN2 over -- having overrepresentation of types which are not likely to be progressing to CIN3 or cancer. So are we talking about CIN3 is the outcome whether as a comparator or using other samples, etc.?

DR. VAN DER POL: I don't think this was restricted to CIN3, was it? I think you just didn't know what to fill in with the table, if it was CIN2. Is that true, I mean, how those would be interpreted, in other words?

DR. FELDBLYUM: Right. Tamara Feldblyum.

This was not restricted because currently CIN2 is used in clinical studies as one of the endpoints. However, if the Panel recommends that that's not appropriate, we certainly are willing to hear that if we're to change our endpoints that are acceptable, so it is up for discussion. It's not a must. We just included it because that's something that has been done before.

DR. FELDMAN: Again, I just want to go as -- go back to the issue from the patient's perspective. So if you see a patient the first time and you do biopsies and they don't have CIN3, you know, let's say they have a small CIN2 but you don't find it and they come back in a year or 6 months, whenever they come back, and this time you find the 30% of CIN3s that were going to progress or develop, you've accomplished something and you've -- you haven't done over-evaluation or management on the CIN2s that were resolving, does that make sense? Maybe I'm --

DR. VAN DER POL: It does, but again, I'm not sure it's part of a study design. It's part of treatment and management, and that's the difference that we're trying to disentangle here is what data do we need --

DR. FELDMAN: I think it has to do with our prediction of risk, so -- it does, it has to go -- so what you're saying is at the first time you come in, if you have this test, what's your prediction, what risk do you have? And we know that some -- some, in particular, of the CIN2s are going to go in one direction, and some are going to go in the other direction, so we don't know where that body goes. So the only way we find out if those become CIN3 is to look at it over time, and then when you license the product, you sort of say it was able to predict the ones that went to CIN3 over time versus it didn't test positive necessarily for the ones that never became -- do you see my point? So I think that there is a value to over time being able to add whether it predicts risk.

DR. WENTZENSEN: I just wanted -- Nicolas Wentzensen.

I wanted to briefly comment on the biopsy procedures. I mean, we have now approved guidelines recommending more biopsies, more than one, for sure, and two or three ideally, and all the trials have done that, as far as I know, and I think the goal is really to have, in women who are referred to colpo to rule out disease and if you take multiple biopsies, we've shown that over years you're -- actually, you don't need to come back every 6 months, every 12 months, and so I think --

DR. FELDMAN: The point is it is really unpleasant to have three or four biopsies, and so those extra third or fourth at that first visit makes it incredibly burdensome on the patients. So doing one or two small biopsies of lesions, which gives you the majority of the yield, is not a big burden on patients. When you add those extra biopsies at the first visit, then you really make it an unpleasant experience for the patient. It is not that unpleasant to come back in a year and have another mild procedure, but it is very unpleasant to have that one, yes, it has a higher yield, but it is extremely unpleasant. I personally never do more than two biopsies; patients don't tolerate it, they don't like it. So it's an unpleasant thing, so you're asking them to have a very unpleasant thing to license a test, and I'm saying I don't think it's necessary because if you just -- what I do is exactly what Barbara said. I just say, well, we missed it, it's probably a small lesion, you know, come back whenever, whenever we tell you, and then if it's something meaningful, we'll find it then. And it's a much less unpleasant experience for the patient. Furthermore, there are some lesions that either resolve or get better and you don't need to know them the first day. If they're really bad, you'll find them, and if not, you can wait.

DR. VAN DER POL: I'm going to try and let some other people get their voices heard here for a minute.

DR. PERKINS: This is Rebecca Perkins.

I was wondering if one of our goals was to try to fill in the blank cells of this table,

and if so, it might be really helpful if the slide could be in a modifiable form, and as we fill them in, then someone could write negative or positive in the boxes. And then I was also -- and that would sort of lead us toward discussing whether CIN2 could be categorized by viral type, so like an HPV 16 positive CIN2 is much more concerning than an HPV 66 positive CIN2, which most people would consider to be like -- or potentially a false positive.

DR. SARAIYA: Hi, this is Mona Saraiya.

I just wanted to ask, clarify, that the CIN2 is going to be already p16 tested because in following LAST guidelines, because that would make it much easier to also fill in the boxes because from some of the population-based studies with this new LAST guideline we're seeing, you know, an increase in the percentage of what we would consider now CIN3s by, like, 20 to 30%. So it would be good to sort of -- I think we'd all be on the same page pretty much if we agree that it was CIN2 p16 or differentiate between p16 positive and p16 negative.

DR. VAN DER POL: So, Joy, if maybe you'd like to comment from the consumer perspective?

MS. HUGICK: I would definitely say that what Dr. Feldman is expressing is important when it comes to the unnecessary burden on the patient, that I would be concerned if you said -- just say there were going to be four biopsies, that you would have a high dropout rate, and that could really impact people's willingness to participate, especially if it's going to be over a period of time. They may not want to come back if they experience a painful and poor experience on the front end; they may not want to continue moving forward.

Was that what you're looking for, or did you need more?

DR. VAN DER POL: Whatever you wanted to add is what I'm looking for.

MS. HUGICK: Yeah, there you go.

DR. THOMSON: Tom Thomson, Chicago.

Can a company bring a new product and request just screening approval or request just approval for a test that is used for clinical decision making because the -- you know, one is a sensitivity issue, the other's a specificity issue, and do we have the same requirements here for both of those requests?

DR. MOSCICKI: Why, a screening test, if it's positive, is clinical decision making, so I don't understand the difference between the two.

DR. THOMSON: So the performance of a screening test relies on its sensitivity, and the performance of a follow-up test would rest on its specificity.

(Off microphone comment.)

DR. THOMSON: Yeah. I used the wrong terminology, I'm sorry. I'm wondering if a new product came along, if they could request an approval for one or the other.

DR. FELDBLYUM: Theoretically, the answer is yes. Tamara Feldblyum for the transcript.

So far all the devices that were approved were approved for screening, either primary screening for HPV or screening as a co-test with cytology, so that's all laid out in the intended use. If somebody came with an intended use to follow up something, it is possible, but we haven't had any manufacturers come in with that, and also, if they were to come up, they would have to follow up to what? It would have to be primary diagnosis determined by histology, something else, and then follow up? So it would have to be a very clear intended use, but yes, theoretically, it's possible.

DR. DARRAGH: Teresa Darragh, UCSF.

A point of clarification for me, maybe, but when we're talking on Question 4A where we talk about the composite molecular comparator, we're talking about the FDA-approved devices and the investigational device. When we're looking at the table and it says a molecular comparator, to me, that is not the investigational device or the FDA-approved

screening devices; it is a tissue-based molecular test like PCR, so there -- the words are a little confusing, and if my understanding is correct, then we really need to make sure that we're talking about the same thing.

DR. VAN DER POL: In the handout that I hope you all got before you got here, there was a little bit more detail in the description of that, and in fact, they were saying that the paradigm might be, if we go with 4B, to have some molecular comparator that was an approved device that would be used, and then in conjunction to that you might add something like p16 or like sequencing from the tissue, which, you know, again would add information to it. So it's actually sort of a combination of both; it's not an either/or.

DR. DARRAGH: So I would suggest that when we're talking about a histologic outcome with a molecular comparator, that it be a tissue-based molecular comparator, not the investigational or the FDA-approved because you're not -- you may not be talking about the same lesion. People can have multiple HPV infections. Just because your, you know, FDA-approved device is positive, that doesn't mean the lesion has that HPV.

DR. VAN DER POL: Patti.

DR. GRAVITT: Yeah, Patti Gravitt.

Could I just come back and maybe try to clarify for myself and hopefully for others as well, is if we really look very specifically at what the study design would entail according to what's been standard moving in the past, and then what we're proposing moving forward, so as I understand it, in the past you had to screen a large number of women because you -- it was a totally untargeted population. You screen a large number of women with good age representation, and then you look at the eventual diagnosis of CIN2/3+ with verification bias adjustment of some detectability. It seems to me that those designs were important in the day because what we're really trying to do is to establish the performance of HPV relative to cytology. Would that be -- yeah.

So now are we at a place where we can all agree that we've established the performance of HPV testing in general, the negative predictive values, the sensitivity, according to what is for sure going to be all -- you know, rife with all kinds of small kind of disagreements, which is just normal clinical practice misdiagnosis, but we've kind of established that.

And so now, from what I understand, the comparator for -- the 4B option is really saying rather than going and screening prospectively, we can essentially say we've already established the clinical performance of an HPV test result by these FDA-approved tests. If I can go into a general screening population and I can compare my investigational test or use with the -- you know, a comparator of FDA-approved diagnostics and I can show non-inferiority and agreement of HPV test positive, that should establish with more efficiency the specificity of the assay.

And if I go into a colposcopy clinic population where I'm going to have a higher yield of CIN3 and cancers, and I can establish non-inferiority of the agreement between the investigational test and the comparator test, then that's establishing essentially similar sensitivity for the detection of hybrid disease. And so, in general, you're going to screen and test a smaller population than if you were prospectively trying to go and find the CIN3, which is going to be rife with consenting women to go through and do prospective follow-up. So I just want to see if that is a good representation of what we're talking about in terms of moving from what was done before to what we're doing now.

DR. VAN DER POL: It is. Yeah.

DR. MASSAD: So, again, this is Leslie Massad.

I would say, yes, but I think we still want to be able to establish specificity and sensitivity to ensure that the device is tuned to minimize overdiagnosis and overtreatment from excessive sensitivity and underdiagnosis and missed cancers on the other end of the

spectrum.

DR. GRAVITT: Would you agree, though, that if you believe that you haven't established sensitivity and specificity for the current FDA-approved devices, that essentially any difference in that is going to be a difference in the, you know, marginal discordant distribution of HPV positive test groups in a routine screen population versus a disease population, so essentially you just look at percent agreement between the test in strata of disease/non-disease because that's -- but because, for example, if you just had somebody who's trying to be super-sensitive, what you'll see is a significant discordants in the general screening population, and that would be an indicator of really poor specificity and that then -- and you can set rules to say if your disagreement is not evenly distributed, you know, on the diagonal, then that's showing nonequivalence in specificity. If it's nonequivalent in the diagonals on the disease population, then that's showing -- it's showing difference in the sensitivity.

DR. MASSAD: So I would say, yes, but once more that that would be looking at a population -- a test that was validated against a different population, and although sensitivity and specificity may be identical, I'd want the reassurance that negative and positive predictive value are retained in a current group.

DR. VAN DER POL: Which is somewhat of a different question but still one that pertains. But, of course, on the other hand, are we going to suggest to the FDA that all tests that are already on the market have to go back because we now have a post-vaccine era and they have to reestablish their performance and they have to reestablish that specificity, and I think that since we're not asking that of existing tests, are we -- again, is there an incremental gain to asking it? I mean, sure, we can find out little bits of information, but overall, from a public health perspective and from a getting market -- getting to the market good quality tests, we don't want to throw out the baby with the

bathwater, and so, you know, we're trying to understand at what point we need to reach for perfection and at what point we need to reach for "this is as good as anything out there," which is still maybe not perfect. And I think --

MR. SPRING: Yeah, Brad Spring with BD.

So a clarifying question, though, because these are qualitative tests, right, so you're getting a yes/no answer, not a load, and so I think your proposal, Dr. Gravitt, was to say if you're going to the referral population, the assumption is they would have the high viral load. The challenge though is, right, if they're CIN3?

DR. GRAVITT: My assumption is you have to do both populations because --

MR. SPRING: Yeah, and then --

DR. GRAVITT: -- you're going to almost inevitably have higher agreement in the disease population --

MR. SPRING: Right.

DR. GRAVITT: -- and maybe lower agreement in the non-disease population --

MR. SPRING: Right.

DR. GRAVITT: -- and the lower agreement may indicate you've got a pretty decent sensitivity, but your specificity has really been declined relative to the comparator, so my assumption is you just have to do both populations because the viral load is inherently different.

MR. SPRING: Yeah, okay.

DR. GRAVITT: And that's --

MR. SPRING: All right, thanks.

DR. FELDMAN: Sarah Feldman.

So I totally agree with what you said, Patti, but I have a question sort of for the group, which is so I understand -- I mean, I actually proposed that, so I totally agree with

what you said, but that implies that we don't think we could do better than the existing test, so the particular population that we really want to know could we do better in are the CIN2s.

So the thing that's very frustrating to people who take care of these patients is there's some subgroup of CIN2s that are going to progress, and there's some subgroup that are going to resolve, and especially among younger women, we'd really like to be able to predict that, and as more tests come out, it would be great if there were tests that could actually improve upon our current predictive testing, both a combination of an HPV test with some sort of other molecular test, I don't know.

DR. VAN DER POL: Sure, but those would show as your discrepants, and that's why if you take the ones that are discrepant between existing FDA tests and a new test, those discrepant ones are the new tests' opportunity to be better or, in actuality, that it's worse. And so if you really dig into those discrepants, and those are the people that you really do histological examinations on, to me, those are the people that really make sense because they could inform practice.

DR. FELDMAN: Well, no. So I'm asking that question, so are you going to require that because the way Pattie presented it, she said sort of you'd look at the sensitivity and the specificity, one in the enriched population and one in the more routine screening population, and I'm asking is part of the requirement that you have to sort of investigate the opportunity to do better.

DR. VAN DER POL: Right. And this was the paradigm I was sort of suggesting when we were on 4 is to go to this model where we really, really focus on those discordants, because it's telling us something, and we need to understand what we're being told.

DR. WENTZENSEN: There's one difference, though. In general, I agree, but I think we would be missing one part of the population and doing that because the women who

come to colpo either had a repeat HPV positive NILM, so they have had a persistent HPV. We don't have the full comparison to the cases to detect it. So, basically, a case would be detected earlier, but they don't go to colpo right away, so we're missing out, like a lower end of the population that the women who are HPV positive for NILM for the first time because they don't go to colpo right away in this population, but we'll get some of those from the general population.

But, I think we could have a slightly less sensitive test to perform equally well because we are shifted to the higher end of the spectrum and still missing out at the low end. So an alternative to doing that would be to not immediately jump into the colpo population but go for an HPV positive population, go into a setting where HPV screening is done with an approved test and take all the HPV positives and a sample of the HPV negatives. That gives us the full population without missing out on any end of the spectrum.

DR. VAN DER POL: And a study that can't be done. It would be quite difficult, right, because to find those people and get them back in and --

DR. WENTZENSEN: No, it would be we're enrolling in a very similar way, and we're doing such a study at the moment with Kaiser --

DR. THOMSON: If I understand Dr. Gravitt correctly, you're really enriching a population, which is one of the earlier questions that we haven't discussed yet, but you're really doing 4A.

DR. GRAVITT: Yeah, in some ways I'm doing a 4A stratified by histologic endpoint. So it's 4A-plus. And I agree that, you know, detailed discrepant analysis would be critical so that you understand whether you were seeing random disagreement, which you will see even if you took the same test, reviewed it twice, it's very common, you know, so you want to just prove that you've got random disagreement, stochastic error, sampling error, things

like that, versus you've got some kind of systematic error which would include, for example, complete types, specific genotyping of the sample to make sure that you don't have a type distribution or a type spectrum problem in one test versus the other.

DR. UNGER: This is Beth Unger.

I completely agree with this sort of two-ended approach, and just getting back to the tissue-based typing, which I agree is really, really important, the only thing I would say is that microdissection is extremely laborious, and for a large study it probably wouldn't be very feasible. And our experience at CDC, which we've done monitoring of CIN2+ lesions with tissue-based typing, is that there's a subset that had multiple types and another group, in their situation where they had multiple types, actually went through various algorithms, compared the algorithms to what they got with microdissection, and there was -- they were going to say what was the best algorithm. It turns out they all performed about the same. So I think we can get the answer we need without microdissection or at least reserve it for a subset, but that's a tiny, little point.

DR. VAN DER POL: Go ahead.

DR. FELDBLYUM: I would just like -- Tamara Feldblyum.

I would like to make a comment that the devices that we're looking at are not approved for typing, so when we're talking about molecular comparator, these are all liquid-based cytology devices. None of the devices for HPV are approved for tissue typing, so please keep that in mind. So what we had in mind when we put this table together, when it says on top molecular comparator, these are FDA-approved HPV devices detecting HPV in liquid-based cytology samples, not tissue typing.

DR. DARRAGH: Teresa Darragh.

Then why is low-risk HPV on that table?

DR. WENTZENSEN: And there's no approved typing, right?

DR. VAN DER POL: Actually, I think when I read through it, it was a combination of using approved molecular assays but also using sequencing and perhaps p16, so there was a typing component to the proposal that's actually in the packet.

DR. FELDBLYUM: True. So the typing component of PCR followed by sequencing is really for the HPV types that are not included in a particular device. If a device comes up with a new output that we haven't approved yet, or if two devices that we're comparing have different types of outputs other than 16, 18, everybody -- pretty much every device has, but anything else, then in that case PCR, some kind of a validated PCR followed by sequencing, for example, would be the molecular comparator.

DR. MASSAD: So, to clarify, my understanding was that tissue-based typing would be for adjudication of histologic endpoints to see if they're high risk or not high risk, not that we would be doing tissue-based typing for everybody.

DR. VAN DER POL: Correct. Other questions or concerns? I mean, I think these topics, the five questions, are also interrelated, which is why we're sort of delving into intended use and we'll come to the -- go back to the beginning and start at Number 1, but I want to make sure that everybody's had a chance to say whatever they'd like to say about this.

And thank you, Mary, for offering to type, and we didn't ask you to after all. Everyone okay, then? No questions about this? So --

DR. MASSAD: I had one more question. That would be 66? This is an opportunity for people to say that they want to include or drop 66? It's on the table.

DR. FELDBLYUM: The reason --

DR. VAN DER POL: With an asterisk.

DR. FELDBLYUM: Tamara Feldblyum.

The reason we included 66, because some devices have been approved already that

detected before it was reclassified, so just to account for the devices that are out there already, we included 66, but we assume that the new devices are not going to include it since it has been reclassified.

DR. VAN DER POL: But so, for example, if you have a new device and it has a negative result when all the comparators have a positive but if that positive the other comparators all picked up is 66, then that doesn't impact the sensitivity estimate for the new device, so that's why that typing would be quite important. And, again, that's why the delving into all the discrepants is going to be really critical.

So I'll just sum this up. So there's not a great deal of clarity about exactly what strategy we think that the FDA should adopt here. I think that's a fair statement. Okay. But I think that there are concerns about, again, burden on the patient and making sure that whatever we're doing makes sense from the patient perspective.

That said, tissue typing directly from the patient tissue is probably critical for understanding some of these outcomes, but it may be that that can be limited to the discrepant analysis, because if we have complete agreement between all the negatives and we have complete agreement between the positives, then those should perform with an NPV and a PPV similar to what the existing tests are already doing.

There was a concern raised that I think is important to note, that, again, although the samples are all collected from patients in the post-vaccination era, having a performance that agrees 100% with a test that was validated in the pre-vaccination era may introduce some differences, and so that's something that maybe should be looked into at some point, and that maybe could be addressed again by typing of some of the samples, even if they're all concordant, to see what kind of impact that's had. Did I miss something?

DR. MOSCICKI: I'm just bringing up the -- at least some prospective follow-up.

DR. VAN DER POL: And so there's some concern, and this goes back to the 4A, 4B

again, that there is some concern that there needs to be prospective follow-up on the patients. And that is not a consensus opinion. I think there's a very -- a fairly strong difference in opinion among the folks at the table about whether that should be part of the clinical trial, and I think we'll talk about that again with the intended use bit, so I think that's not quite done with discussion yet.

DR. BURK: You didn't mention what the histologic diagnosis for your molecular comparator would be. Is it CIN2 and CIN3, or is it just CIN3?

DR. VAN DER POL: At this point, CIN2, I think, is important to keep on the table. There was enough consensus around the table that we don't understand how best to diagnose those CIN2 lesions that do lead to cancer, and so having a diagnostic test that's picking up or differentiating between those would be useful. Therefore, looking at CIN2 as an outcome is probably still worthwhile.

DR. DARRAGH: Teresa Darragh, UCSF.

Again, in my mind, the CIN2s need to be supported by at least a positive p16, if not other tests that come in, in the future that are better predictors of transformation, because CIN2 is a wastebasket and poorly reproducible by pathologists on H&E stain alone, and it includes a lot of things that other pathologists would call CIN1 and also includes things that are just benign. So at least make that gold standard a little bit more gold if CIN2 is included in the outcome.

DR. VAN DER POL: Any other comments on this topic that I didn't catch?

DR. KINNEY: Walter Kinney.

I don't think it's widely appreciated how unrelated CIN2 is to cancer risk. I see Dr. Schiffman is no longer here with us, but he published -- perhaps I'm missing him.

(Off microphone comment.)

DR. KINNEY: I see. Dr. Schiffman published 1600-some untreated CIN2. At a median

follow-up of 48 months, there were three cancers, one of which was related to gross labs and follow-up compliance, and the other two of which were arguably missed CIN3 at colpo. And if, in fact, the risks are that low in the Tainio meta-analysis that she showed would support that, then why you would want to use CIN2 as an endpoint is unclear to me.

DR. VAN DER POL: So there's some disagreement about whether CIN2 should be included as an output. And, again, I think we want to be really careful here to disentangle patient management and a diagnostic test, and so I keep coming back to that because that's the hat that I wear the most often. So if we're trying to look at how does this test perform, particularly if the question is how does this test perform compared to tests that are currently available and moving our ability forward, that's where we have to make sure we're keeping our focus.

MR. SPRING: Just -- sorry, Brad Spring.

Just to add, I think, as Dr. Darragh said, the importance of CIN2 is around the methodology, right, and use of p16, and I think some of the previous publications in other studies didn't have the use of p16, if I'm not correct. And so I think the advancement of some of the classification methodology may change that paradigm, and I think, FDA, you've even acknowledged some of that in recent guidance, too, right, about the evolution of that.

DR. MOSCICKI: Barbara Moscicki.

But I do want to clarify that we seem to always be falling back to say we're only looking for prevalent disease, and right now we actually look at like 3-year risk or 5-year risk. So this is where I still think the endpoint is CIN3, but you would continue to prospectively follow a CIN2 because I want to know what that person's 3-year risk is for CIN3.

DR. PERKINS: Rebecca Perkins.

And I think that's especially important when we're -- and I know I'm about to leap to

Question 2, I think, but all of these tests are probably going to be approved for primary screening in a general population, and right now we're talking about following them for 3 years, but the reality is the guidelines say 5 years. So if we have only one -- if we have a new test that has only one point of data that's immediate risk and we're saying that it's going to be okay for 5 years in the prospective data, I think that's a little bit of a leap potentially.

DR. VAN DER POL: But if the test is performing identically to tests on the market, there is no leap. I mean, that's -- again, disentangling that diagnostic time point from what clinical follow-up does, you know, they're very different things. And so if you were using what's available on the market and you have a new thing that performs exactly the same, are you any worse off?

DR. PERKINS: Right, I really think exactly the same -- you get to a power issue because exactly the same when applied to the entire population of the United States versus exactly the same at a single time point in a test that may not be identically reproducible in a small sample size may not give us as much confidence of exactly the same.

DR. VAN DER POL: Sure, fair point.

Okay, I'm going to wrap this up for this topic. So I think that -- I think we've covered all the points in the summary, so I think everybody's got everything here.

So let's move on and pretend like we're at the very beginning. And let's do Question number 1 now, which we're going to talk about proposals for ways to actually enrich or supplement the populations that are in these studies to help these studies be able to be accomplished in a less burdensome way. So here the proposal is do we recommend a -- one or a combination of the following three proposals to increase the number of women positive for CIN3 and/or high-risk HPV in clinical studies. So supplementing from referral clinics, like colpo clinics which we've already been talking about, utilizing archived

specimens, and capping vaccinated populations. So I think it would be beneficial, actually, to talk about each of these in turn, but I think we've already touched on the first one because as I said these are very overlapping. It sounds as if there's already fair consensus that we could supplement from referral clinics but that it has to be managed well.

DR. MOSCICKI: I'd also like to just comment on that there are two different screening populations and one that has been screened really well, private practices, etc., and then you have the women who come in who have never been screened before or are drastically under-screened, and that's really the majority, if you look throughout the United States, but obviously they tend not to be in anybody's clinical trials because follow-up can be difficult because they often are immigrant populations.

DR. VAN DER POL: Good point.

Other people have thoughts about -- I'm going to stick on the referral for colpo. Do you think this is a good idea? Are there things, caveats, that you would like to put into that? Do you see risks associated with that, things that need to be managed up front?

DR. WENTZENSEN: Nicolas Wentzensen.

I think it's really important to define how referral happens in these clinics because we still have all kinds of different practices happening in the U.S., we have cytology screening in some settings, and I think that we would want an HPV screened population, either co-testing or HPV alone, to really go to these referral clinics. If we only take any referral population, there could be very different populations in there, so I think it is important to define. And there's also a lot of management happening at referral clinics, so women who have been tested multiple times who are in a post-colpo setting who are retested many times, so I think it would be very important to define what referral population we're talking about.

MR. SPRING: And just maybe a question for the Panel around what would be

potentially the ratio of, say, referral or an enriched population, and it may be dependent upon the intended use, of course, which is something to consider, too. Is it 100 percent? Probably not based on the discussion, but what would be that percentage maybe aligned with a particular intended use for a new test, not necessarily for a modified test.

DR. MASSAD: This is Leslie Massad.

I would think you would want similar outcomes of interest, which means that you'll have to have a fairly large screening population to find those smaller proportions of outcomes of interest.

DR. VAN DER POL: Well, if you go to colpo clinics, you're going to enrich for that, which is kind of part of the whole conversation.

Dr. Lawson, I believe, has a question.

DR. LAWSON: Yeah, I just had one concern. If you're looking at a referral clinic, say, versus a lower volume source of colposcopic examinations, shouldn't there be some form of, you know, accepted recordkeeping that would provide you with the kinds of information that would be useful in this setting?

DR. VAN DER POL: So as far as standardized datasets that people would be required to pull from an EMR before they could consider a person to be eligible, is that what you're asking?

DR. PERKINS: Rebecca Perkins.

Could I suggest a fourth potential option, which would be targeting underinsured women who may have -- who may not have gone for screening and treatment due to financial concerns and then enrolling in a trial where they're getting excellent care because they're getting a standard test against a new test. This might improve healthcare for the most underserved women, have a very high-risk population and yet it's also a screening population. So it would improve public health while also testing the new device.

DR. VAN DER POL: What types of clinics do you think would participate in a study like that?

DR. PERKINS: Planned Parenthood.

DR. VAN DER POL: Planned Parenthood is usually already a participant in most of these trials, so that wouldn't be any different.

DR. MOSCICKI: Department of Public Health clinics are the most frequent ones, certainly like in California.

DR. VAN DER POL: I'm just playing devil's advocate here because -- so at my clinic, when I was in Indianapolis, we did not perform any sort of cervical cancer screening at the STD clinic, and in fact, there were a few OB/GYN clinics that were public health clinics people could go to, but that was fairly limited. In Jefferson County now where I'm in, in Alabama, we actually do cervical cancer screening at the STD clinic, and the only reason I throw that out there is (a) it's the population I'm familiar with, but (b), they are almost always very underserved populations.

And so that's one of the ways I always see to enrich that. But then if you're talking about having a dataset of some sort, it's a pretty transient changeable population, and so again, follow-up can be a challenge but not worth not chasing, but I'm just saying if we're looking at the logistics of -- part of the question here is how do we make this less burdensome, and so I think more options, more clinical options, are a way to make things less burdensome, but we also need to think is this something where a lot of the health department clinics anymore don't like to have industry studies in them because they feel like that's an abuse of their patient population. So there's lots of things to weigh into the balance, all I'm saying.

DR. SARAIYA: I was just going to comment, the CDC's breast and cervical cancer screening program is another option. We actually try to recruit women who are never or

rarely screened as part of each state's population, so there is an extensive amount of effort that's put into recruiting those women, so that might be an option. But, again, I see your point that it's public health clinic. And then just to comment on the STD clinics, yeah, it's -- we've done a survey of STD clinics, and it looks like maybe 30, 40 -- 20 to 30% of them do screening, cervical cancer screening. A lot of them choose not to do screening because of the concern about follow-up.

This is Mona Saraiya.

DR. VAN DER POL: Comments about -- can I just see like nods because I think we're all in agreement that referral in some form or fashion makes a lot of sense to enrich in the population. Does anybody see any risks associated with that, that we've not discussed?

(No response.)

DR. VAN DER POL: So now let's move on to using archived specimens and what comments the Panel may have about the potential benefits but also the potential risks, and I'll just start with archived specimens, even if people say that like, for example, if you're using PreservCyt, you know, and you can store it and people say they're freezable, but of course, it doesn't freeze and they evaporate, and so then concentrations change. So there are issues because of the nature of this particular medium that I think it serves us to be cognizant of those issues. So while, in general, I'm in favor of that kind of approach, I'm not 100 percent sure it's workable in this case, but I'd be welcome to hear other people's thoughts.

DR. DARRAGH: Teresa Darragh, UCSF.

I think looking at what we would use these for is important, so using an enriched population either by, you know, a colpo or leap referral population, the archived samples of women with proven disease, would be to look at the true positives when our screening population, we're going to be more concerned with looking at the true negatives, and you'll

get different information out of each population and it's -- just, to me, it seems important to keep those two ideas separate.

DR. VAN DER POL: And are you raising that in the context of archived specimens as well as the earlier conversations?

DR. DARRAGH: In archived specimens in that, you know, with those you'll have the endpoints for those with disease. The one question with that, that you raise is also about the degradation of specimen over time.

Nico, didn't you do that with a Costa Rica population and looked some degradation and --

DR. WENTZENSEN: Yeah, I mean, we've -- I mean, that specific example was dual stain, which is a different type of assay that we're not discussing, but I think that there is concern over time, even if you freeze specimens, that you may have a decrement, and particularly if we're talking about DNA versus RNA, I think there could be issues for non-DNA tests that are -- that would have to be worked out and really demonstrated that they were from the archived specimen, so I have some concerns.

At the same time, I think there's a situation where we have archived specimens from large like PMA trials that were conducted previously that are well-characterized and now somebody wants to go back and look for an updated version of a test, and again there, I think, there is an argument for using these archived specimens rather than redoing the whole system. So I think there's different situations. There's the one where you have unknown origin specimens that I would be very dubious about, but then there's the large trial that has frozen aliquots, and I think those would be very different.

DR. MOSCICKI: Barbara Moscicki.

I would also just to like to comment, not all aliquots are equal. As somebody who's done it, I find a positive-negative-positive in my own aliquots, so we think we're distributing

HPV equally, but I don't think we always are.

DR. KINNEY: Walter Kinney.

One of the reasons to struggle through the specimen problems associated with saved specimens is that it gives you a view of longitudinal changes over time, and that's priceless in large quantities. Dr. Schiffman and company have not only specimens over time but the clinical attributes for those individuals, and that's something that it takes a great deal of resources to assemble outside of that setting.

DR. UNGER: I'm basically agreeing with the other comments. I think the archived specimens can be really valuable, but I would be very concerned about having them be the only source because of differences in preservation and that can affect even -- even the different devices do different kinds of processing of the specimen, so you kind of lose that when you've stored something for quite some time, and that could be a really key step that's very hard to monitor, but it is really important. Again, it gets to sort of like what the -- if we do the analytic performance regardless of how you got there, I do think the archived material is very, very helpful.

DR. ALEXANDER: Barbara Alexander.

I've done a lot of this with fungal diagnostics, and it's critically important to understand the integrity of the sample, like how was it processed, was it DNA free from contamination when it was put into the tube, what kind of freeze thaw cycles has it been through. Also, for RNA-based assays, which I'm assuming will be coming to this field, these -- the RNA is very unstable, and so without stabilizers of some sort, you're probably not going to have RNA that's detectable. And then, finally, I think your issue, Barbara, around even if you have the same specimen, different aliquots of that same specimen may not test equitably, right? And so it would be also important to have, from the same aliquot, to have your comparator tested against the new device from the same aliquot. And you also can't

use historical testing results against your new device from a sample. So if your sample is there from a PMA application or whatever and it was tested against a device that's currently on the market, you still can't use that old result to compare your new result, too.

DR. WENTZENSEN: Sorry, I forgot one point. If you have archived specimens from a previous study where you may have excluded specimens because of inadequate results, you may have a biased population, so you have to be really careful that you, again, reflect the whole population.

DR. BURK: Robbie Burk.

I mean, I would agree with what Barbara said, that the efficiency of using archived specimens gives you longitudinal data, and though they do have caveats, you know, simultaneously testing with the old device and the new device, you know, it's certainly worthwhile just do the efficiency of the study design of the original materials.

DR. ALEXANDER: So my point would be that I think the archived specimens are critically important, especially in a disease state like this where it's not -- you know, there's not a lot of it around, like we have with invasive astrogliosis as an example, but it would be important to use not only archived specimens, be smart how we're using them, but also then have a certain percentage of the specimens be prospectively fresh-collected specimens and tested.

DR. GRAVITT: Patti Gravitt.

I have just one kind of logistic question with regard to the use of archived specimens. How would the ownership be determined? And then if someone was interested in going back and testing a new investigative device using the archives, who makes the decision of whether that's an efficient or appropriate use of a resource that somebody's put a lot of money into establishing, and would that be an equitable decision?

DR. VAN DER POL: I mean, I think that's a great point to consider because if it's

something that -- for example, if it was something that was at NCI, then tax dollars paid for that, and so, you know, if you gave them out to four companies and the fifth company comes and you say we're out, then you start having some real headaches. So I mean, I think those are real logistic issues that people would really have to take into account.

DR. MOSCICKI: One other caveat for the archived specimens would be if you do have discrepant results and then you have to now somehow access biopsies and there may be no biopsy left or you're looking somewhere else, I think it would be really -- I mean, what do you do with the results if they're discrepant?

DR. VAN DER POL: Other thoughts about archived specimens?

(No response.)

DR. VAN DER POL: So then the last topic in this particular question is what do people think about capping the vaccinated population, remembering that this may even be a temporary question, because if the vaccinated population in the U.S. as a whole becomes large enough, then you probably can't even approach this strategy. But for the time being, while this may still be an option, do people think that that impacts the evaluation of test performance?

DR. MOSCICKI: I have a whole question around HPV vaccination. It's how do we even document it? Are you going to ask women to come with their vaccination records, show that they had one, they had two, they had three, and at what age, otherwise you won't let them even enter a study? There's going to be so many women who do not carry their vaccination records around, I can guarantee that and, you know -- and asking them at what age they got vaccinated, well, and which vaccine.

(Off microphone comment.)

DR. MOSCICKI: Yeah. So I mean, just a whole issue of including vaccination, women, I think, is going to be so messy that it -- you know, but I think we have to figure it out. I'm

not saying you can't, but you're going to have to mandate some type of documentation, but then do you also document un-vaccinations? Show me you haven't been vaccinated.

DR. VAN DER POL: Well, I'll just tell you, if we use my son as an example, if you ask him what he's vaccinated for, and he's 31, and he would go, I don't know, ask my mom.

DR. MOSCICKI: So that has to be part of the design, right, mothers --

DR. VAN DER POL: To call mother. But no, it's a valid point, and that's the first thing that I wrote down when I was looking at it because I'm like really, you know. And, you know, but I mean, people don't know. My husband is the same way, I mean, so I don't know if that's just my little *n* of 2, but people often just don't know if they're vaccinated or they don't know, especially now if you talk about was it the bivalent, the quadrivalent, the 9-valent, people aren't going to know.

DR. PERKINS: We did a study on this looking at actually parental recall of getting their kids vaccinated, you know, within a few years of vaccination, and it was not very reliable.

DR. ALEXANDER: Barbara Alexander.

I have a question and -- because I'm kind of new to this area of HPV, but bringing in my day-to-day activity of caring for transplant immunocompromised patients, has there been any thought of or research into these patients who are iatrogenically immunosuppressed, and we know that they have an increased risk of other pathogen-associated cancers, like CAPCs, PTLD with EBV and so -- and we know from studies that the rate of cervical or HPV-related infections is increased in this -- transplant population, for instance. And so when I looked at this question, trying to figure out if there was a way to supplement patients that might be a high risk for progressing more quickly to a CIN2 or a CIN3 from screening, you know, from time of transplant to diagnosis of CIN3, is there any role of -- I'm trying to capitalize on that population somehow.

DR. MOSCICKI: I would just like to comment that, in fact, we just did a big review that hopefully will be out soon, in the lower genital tract, on basically iatrogenically immunosuppressed individuals. The sad part is there's just so little data out there, so -- and it really varies from what autoimmune disease do they have to transplant, and then we're using drugs we never used, you know, 10 years ago when all that data came out. So it's really a mess, and if I was a company, I would not go after that group at this point since it's just such a mixed bag.

DR. LAWSON: Herschel Lawson.

I'd like to be really optimistic and think that this is going to be a really short thing and that we probably shouldn't even discuss it any further. I really don't think that there's much advantage to the use of limiting the population, and I think that we should push in the direction of having even more vaccinated.

DR. DARRAGH: Teresa Darragh.

I agree with that wholeheartedly. I actually think for those of us who are in medicine, it is a bad message to say I'm looking for unvaccinated women, period.

(Laughter.)

DR. VAN DER POL: Can you use the microphone?

DR. PERKINS: I just said unless the second half of the message was I'm looking for unvaccinated women because I'm worried they have cancer and I need to give them my test.

DR. VAN DER POL: See, there's a positive way to spin it. I think the other thing is that my recommendation on this personally would be to look to Australia. Australia has, you know, just excellent vaccine coverage, and they actually are collecting data, they've been collecting data for years, and Australia and the Netherlands are the two countries that I would say, you know, really -- both have population-based health services, so people don't

fall out because of lack of insurance, and both have records that are consistent for people regardless of where they're seen in the country.

So they are doing these types of analyses, and so I think we will know in the next few years, you know, what they're seeing in terms of whether or not the proportions must change in terms of which HPVs are causing cancers, but the question really is here are cervical cancer rates dropping, and it looks so far as if that's going to be the case. So I'm not really sure that having a diagnostic test that doesn't pick up these other ones, that it would really impact what we're doing in terms of clinical care right now, but I do think that there are probably data out there that could be examined, but I agree that this is, knock on wood, a short-term question. That was one of the things posed, and so we can provide that as our consensus agreement here.

DR. DARRAGH: Teresa Darragh.

Just a question. Do all of the trials that go before the FDA have to be done on U.S.-based populations?

DR. FELDBLYUM: It is preferable, but it is not a requirement, so especially for infections that do not occur in the United States or that are very rare, we do accept foreign data. Actually, FDA published guidance or an SOP on use of foreign data in clinical studies, so partial foreign data is acceptable, but we certainly prefer U.S. data because that's where the device is going to be used. So you want to use, in your clinical study, the intended use population to show how the device works in that population. That's why we emphasize preference for U.S. data, but we're certainly open to discussions if somebody has other suggestions.

DR. VAN DER POL: And it's unclear whether you would want to do a new diagnostic evaluation in someplace like Australia or you'd rather do it in some place like Kenya, right, because it depends on what question you're trying to ask.

DR. BURK: Just a comment about capping the vaccine -- so in the United States, once we're hitting 70, 80%, you can just do a cohort and assume that either they've been vaccinated or they're getting kind of herd immunity protection. So that would be sufficient, in my view, for -- you know, if we want to sample that group.

DR. UNGER: So I guess I maybe misunderstood, but I thought the idea of capping the vaccinated population was to enrich for endpoints. I thought that was the question.

DR. VAN DER POL: Right, by reducing the number of people who are vaccinated in your trial.

DR. UNGER: Right, right.

DR. VAN DER POL: Right. That is right.

DR. UNGER: Yeah.

DR. VAN DER POL: Mr. Spring.

MR. SPRING: Yeah, so I agree that the scope was around trying to get CIN3s higher, but the vaccination question may be important when you look at genotyping, right. So as prevalence of 16 and 18 supposedly goes down, and then you want assays to look at other genotypes, you may need to understand whether someone has been vaccinated or not because now you're going to be talking potentially enriching a population of other genotypes and how do you get them into the trial. So it's probably not for this particular question, but it may be important for more future-looking indications.

DR. WENTZENSEN: Related to that --

MS. HUGICK: I was just going to -- this is Joy Hugick.

I was just going to piggyback on that and say maybe it shouldn't be a requirement, so maybe you shouldn't use it to exclude people from the study, but perhaps it should at least be a question asked up front.

DR. VAN DER POL: I think that's pretty much the standard or -- and if it's not, I agree,

it certainly should be.

DR. WENTZENSEN: Nicolas Wentzensen.

I just wanted to address the comment, no matter what the vaccination says, given a certain type, the risk is pretty much the same, because like the vaccine either prevents the type from being there or the vaccination didn't happen or there was a whatever it is, the risk of a type is the same no matter if the vaccination --

MR. SPRING: It was just getting the numbers of a certain genotype, that was just more of the comment, but yeah, I agree, the risk is no different.

DR. SARAIYA: Just a comment. Mona Saraiya.

From a public health system, every state is sort of increasing its adolescent immunization registries, so these are verified records that companies may be able to take advantage of. Somebody was asking whether it has to be done in the U.S. or not, just to want people -- remind people that there are several places in the United States that are underserved, under-screened populations like the U.S. Pacific islands, the Caribbean, and there's also a high vaccination rate in some of these islands as well.

DR. MOSCICKI: As someone who vaccinates adolescents, I would say 40% are in those registries, and the other 60%, their records are in somebody else's private office.

DR. VAN DER POL: Other thoughts about capping vaccination as a study strategy?

DR. MOSCICKI: Actually, Rebecca commented just get the older women in. We know they haven't been vaccinated.

DR. VAN DER POL: Except we have to have a good distribution of age in the clinical trial population. There, I quoted something.

DR. PERKINS: I was just thinking that if vaccination proceeds as we hope, then the women left most at risk for cervical cancer will be those who just missed out on getting vaccinated but also weren't screened very frequently throughout their lives, so that may be

a population that is sort of accidently enriched for disease.

DR. VAN DER POL: Yeah, I mean, certainly worth looking at when we're trying enriching strategies, right?

Okay, so I'm going to sum up, and nod when I'm right and shake your head when I'm not. So the Panel would recommend supplementing from referral clinics; however, the Panel thinks that this needs to be well controlled and well described so that we understand which referral clinics, who they were referred by, what they were referred for, how many times they've been seen, and a number of variables of a similar nature, so we would recommend that that be very well controlled. The Panel agrees that archived specimens are a valuable resource that may provide a benefit to studies; however, there are concerns about how they are stored, what they are stored in, who has access, who owns the samples, and given those kinds of concerns, it's unlikely that that's going to be a strategy that most people can use going forward. Is that pretty much a summary of what we -- okay, I'm seeing pretty much yes.

And then, finally, as far as capping the vaccinated population in general, I would say that the Panel recommends that this is probably not a worthwhile strategy because it's too difficult to really identify who those patients are with any accuracy because it would involve knowing not only what you were vaccinated with but when and how many doses, and the level of that kind of recall is somewhat difficult, and hopefully that will become a moot question, and so we think that it's probably not necessarily a good enrichment strategy.

However, there may be other enrichment strategies that people could look at. Some that the Panel recommended were considering reaching out to underserved women, women who have not previously or not historically been screened as frequently as perhaps they should have been, and also reaching out to older women who missed sort of the vaccination of -- and in terms of a birth cohort, who missed by age the vaccination

eligibility, and that would not be your approach for the entire study population but as an enrichment strategy. Did I leave anything out?

(No response.)

DR. VAN DER POL: Okay, then you've all earned a break.

(Laughter.)

DR. VAN DER POL: Okay. While you're outside, please don't discuss what we're discussing in here with other Panel members or anyone else, and I will see you back in 15 minutes.

(Off the record at 2:31 p.m.)

(On the record at 2:55 p.m.)

DR. VAN DER POL: Okay, welcome back. We've done the vast majority of the day's work, so thank you and yay. And I assume that we're going to get out of here early and -- but make sure that if there's something that you want to say, that you take the opportunity to say it and don't let me move past you.

So we're going to start back up with Panel Question 2, which is talking about with the clinical studies that have been performed in the past, the routine method has been that people with NILM or H -- high-risk HPV double negative people and ASC-US and HPV high-risk double negative people enrolled in clinical studies have been still called in for colposcopy. And so the question here is do the benefits of colposcopy referral in these cases outweigh the risks associated with the procedure and the potential overtreatment? So they'd like us to discuss each of those populations separately.

Remember that with the NILM double negatives, those people, only a sample is pulled in for colposcopy, not the entire population.

DR. MOSCICKI: I would just like to ask if a company was going after just primary HPV testing, do they even have to bother with cytology if they're just going for -- and they didn't

want to do co-testing?

DR. VAN DER POL: Which is an interesting and delightful question.

DR. FELDBLYUM: Tamara Feldblyum.

So far we did not have anyone come in with an application just for primary testing. It actually historically started the other way around. The companies got approval first for adjunct testing, and then they came in with a primary testing indication. So, historically, this was the method that was used. If somebody comes in with a primary only, we'll have to certainly consider different study designs, but we just didn't have such an application so far.

DR. VAN DER POL: And this will go back to Question 3, which we'll get to in a moment, which is the intended use question, so keep that sort of in the back of your mind as well. And then so --

DR. MOSCICKI: I'm Barbara Moscicki.

I do not feel this day and this time, that it's important to do this negative verification bias and that they do not need rescreening in 5 years for the first NILM and high-risk HPV double negative.

DR. VAN DER POL: That's not the question here. The question is in the clinical study, a random sample of these people are pulled in for colposcopy.

DR. MOSCICKI: I am saying no.

DR. KINNEY: I would agree.

UNIDENTIFIED SPEAKER: I would agree.

UNIDENTIFIED SPEAKER: Agree.

DR. VAN DER POL: Can we see hands real quick? This conversation's over.

(Show of hands.)

UNIDENTIFIED SPEAKER: Right answer.

DR. VAN DER POL: I really love consensus, yay. Okay, so that was the NILM. Does everyone feel similarly about the ASC-US because the ASC-US is a different population so -- that think that colposcopy is not warranted in those populations, HPV double negative ASC-US.

UNIDENTIFIED SPEAKER: ASC-US HPV, right.

DR. VAN DER POL: Okay. There is not going to be a lot of conversation. Is there a question at the end?

DR. FELDBLYUM: No, it seemed like one person didn't raise their hand. If there's somebody who would like to provide a comment?

DR. THOMSON: I should have, I'm sorry.

(Laughter.)

DR. VAN DER POL: Inadvertent absentia.

DR. FELDBLYUM: I just want to make sure that everyone is heard.

DR. DODD: With regards to the double negative, this is Lori Dodd, I just want to ask you a clarifying question about how the double negative is being defined. Is it any of the predicate tests or --

DR. VAN DER POL: Well, that's an interesting question, and again, this relates back to current -- I mean, past study designs and in fact, in general, these had one molecular test that they were compared to, and cytology. So if you go to a different paradigm where you're now using multiple molecular tests, that's where those double negatives are going to come from. But historically, that double negative actually included the investigational device, which is not very comfortable for most of us, but we're going to overlook that and talk about going forward with different devices, not the investigational device because I don't think we want to make any clinical decisions based on those. Does everybody agree to that as well?

DR. DARRAGH: This is Teresa Darragh.

In the --

DR. VAN DER POL: Can you pull the microphone towards you?

DR. DARRAGH: In the paperwork, it says that double negatives are defined as negative on both the investigational and the FDA approved.

DR. VAN DER POL: Right, that's what I was trying to say. That was the older paradigm, and what we're saying is that going forward that can't be the case, really, that the investigational device shouldn't be part of what's counted in the double negatives.

DR. PERKINS: So they'd then be basically triple negatives? They'd be negative in the investigational device and negative on two FDA-approved devices?

DR. VAN DER POL: Yes, but I wouldn't even look at it that way because the investigational device is not part of the conversation, really, so you would still call them double negatives but just a different definition of negatives.

DR. DARRAGH: But why do you need two negatives on the FDA-approved device given the exquisite negative predictive value of the FDA-approved HPV test?

DR. VAN DER POL: Because when you're doing these clinical evaluations, a general rule with the molecular test is because you can have something that was introduced by an instrument spitting or you can have something introduced by a gloved hand or just all kinds of things where you can either have sample contamination or you can have amplicon contamination, and so just as a 100% -- or not 100%, but as an attempt to be as detailed as possible, if you have multiple comparators and so those double negatives by multiple comparators, everybody is 100%, sure those people are negative for HPV. Whether they're negative for cancer is a different issue.

DR. PERKINS: Sorry, just to clarify. So someone was negative on -- I know you said we weren't talking about the investigational test, but if the investigational tests were

positive but the other two were negative, that would be discrepant, and then they would get evaluated, correct?

DR. VAN DER POL: Correct.

DR. PERKINS: So these are people who are negative --

DR. VAN DER POL: Correct.

DR. WENTZENSEN: I'm not --

DR. VAN DER POL: Go ahead because there's -- I can see questions on faces.

DR. WENTZENSEN: Yeah, I'm not -- because with what Rebecca just said, I mean, the investigational test is part of the decision, I mean, if they're positive in that.

DR. VAN DER POL: The investigational test can't be part of the decision for patient management.

DR. WENTZENSEN: Oh, yeah. Yes.

DR. VAN DER POL: So, no, you wouldn't drive colposcopy based on an investigational positive when your two FDA-cleared tests were both negative.

DR. WENTZENSEN: Didn't we just say the opposite?

DR. VAN DER POL: You would look at them as a discrepant, right, and so what kind of follow-up you would do if you're -- right, as part of the trial, right. But I'm talking in terms of just routine clinical you wouldn't.

DR. DARRAGH: So can we have the FDA clarify what they meant in Question Number 2 about double negatives, if they meant two FDA-approved negatives or the one FDA negative and investigational negative?

DR. FELDBLYUM: So the question was based on current or past practice of how the clinical studies were designed, and in that case, there was an investigational device negative and an FDA-approved device negative. Going forward, we'll have to see what kind of clinical study design, based on what the Panel is recommending and what we can

implement, it may change. But so far, the definition was true that that was one investigational negative and one approved negative.

DR. VAN DER POL: But while we're on this topic, I think maybe the kind of underlying topic that maybe goes with this is for clinical studies, is cytology a metric that should be included, because that's part of what this is asking is that when you have NILM and multiple molecular test negative, however you came to those molecular results, or you have ASC-US and those -- because if you were doing just routine Pap and you got an NILM, you're not going to refer that person to colposcopy. So should you be doing that as part of a study just to fulfill study requirements is kind of the question here, but I think the adjunct question could potentially be do we think that companies should be using cytology as part of the classification scheme?

DR. MOSCICKI: If it's co-testing, you have to.

DR. VAN DER POL: And so in those cases, then, we're still all agreed that colposcopy should not be required for these women that have multiple negative molecular tests. If it's a discrepant result and we're doing discrepant analysis, that's a different issue. Okay.

Okay, I think I can sum that one up pretty quickly. Okay. So the Panel had -- I think it may be worthwhile because I think now I'm confused, which is not difficult, but it may be worth clarifying kind of very, very specifically, you know, who is being consented to what in this alternative study design? Is it looking at an enriched colposcopy population against a general screening population and consenting them in to have multiple -- or to have a sample collected that then is divided into multiple tests but their routine standard of care is what's driving the detection and the histologic diagnosis, or are they being consented to actually be followed under study protocol for a colposcopic referral?

So I'm a little unclear exactly of what you're consenting women to do in this trial versus what you were doing in the past. So in the -- well, I guess that's part of this, too, but

so in the past when you were consenting a woman into one of these studies, they would have the specimens collected for the molecular test, the one under investigation, and if there was already one available, and then they would be managed according to cytology, and so they would have that algorithm to say if you were ASC-US but HPV positive by a molecular test, you could go to colpo. If you were at -- if you were CIN2 or 3, obviously you would go to colpo. But if you were ASC-US and both molecular test negative, you would still go to colpo because you were ASC-US, or at least a portion of the study people would go into colpo.

And if you're NILM, which nobody would go to colpo normally, but even if you were NILM and both of the molecular tests were positive and you were consented into a protocol where you had some probability of being selected to go to colposcopy, and so the question here is do the risks of that outweigh the benefits?

DR. GRAVITT: Yes, I understand that with regard -- and I think we're all in consensus about not necessarily referring the known HPV negatives, but my question is who's dictating the management? Is it just routine clinical management and you're just comparing test results in the given populations, or is the study design trial dictating here is who I am going -- standardizing who's going to be followed versus you're going to retrospectively just analyze your comparison data --

DR. VAN DER POL: I think the study design has to dictate who gets which procedures because that has to be part of what they're controlling data collection on. But then if the study doesn't think that -- if a person's not referred to colposcopy because of the study design but the clinician thinks that person needs to be referred, then of course that would happen as part of standard of care.

MR. SPRING: Brad Spring of BD.

Just to remind folks, the current study design, though, is kind of this "or" algorithm.

We do look at, you know, ASC-US or NILM. So if a patient comes into the study, they are NILM and the FDA-approved test is negative but the investigational test is positive, they are referred to colpo, right? Does that make sense? But it sounds like we're debating now they should -- the investigational test should not be used, then, to refer them. Is that true?

UNIDENTIFIED SPEAKER: But that's not the question here.

MR. SPRING: No, I just --

(Off microphone comment.)

DR. VAN DER POL: Right, but would the trial -- so would the trial write that in, is that if their investigational device was positive, that the person would be sent for colpo and the trial would probably write that in because you'd probably want to know, right? But that's not the negative-negative group that is of concern.

Okay, so we're going to move on to Panel Question 3, which is Panel Question 5 in actuality. And this is the one that I think maybe, you know, a robust discussion because I think that one of the things that we've been kind of talking about really all afternoon has been intended use, and we're talking about whether something is a screening tool and whether it's the negative predictive value that we care about or whether it's a diagnostic and we care about the sensitivity, and so I think that this is really probably the crux of the entire afternoon for all of us.

So the question here is do the benefits outweigh the risk of consolidating the indications to encompass one general screening population and removing references to specific triage tests and clinical actions? So making the intended use statement to be way more generic, and there are issues to be thought about including age and some of the other things that come to play here and when you're trying to compare sort of apples to oranges with old tests and new tests, so --

DR. WENTZENSEN: I have a question to FDA -- Nicolas Wentzensen -- related to that.

And the question is if the test is evaluated, it still will go through a certain process of a screening test and a triage that follows, and then like there is like some strategy that is being used and the approval will be based on that strategy and then are we then -- are you proposing to then make the language more general to move away from the strategy, or are you -- so is the evaluation different? I'm trying to understand that difference here because --

DR. ZARITSKY: Yeah, so if the language is more general, then we would be moving away from that strategy of within the approval process evaluating the device coupled with the triage test.

DR. WENTZENSEN: The problem I see is that -- I mean, we see more and more, as we evaluate more triage strategies, that we have to evaluate screening and triage together because only the combined strategy gives you let's see who goes to colposcopy, which disease is found at what time point, and you basically have to consider both steps to really be able to evaluate a test.

One example is if there's an approval for a test to send HPV 16 positive women to colpo but we have now a great triage strategy that can better triage HPV 16 positive women and we are sending some of them not to colpo, is that a possibility, because previously that was kind of a problem when 16 had to go to colpo. Can we move away from that? If the first candidate like includes 16 referral to colpo but now we have a new triage strategy, can we --

DR. ZARITSKY: Yeah, so that's the exact reason why we kind of want to move away -- or why we're offering a proposal to move away because these new triage strategies can affect previous strategies before. So I do want to point out that the new -- if there is a new triage strategy, that will be evaluated by the FDA with some representative HPV test. So that will still happen. It's just that for HPV device manufacturers, if this new proposal for an

indications for use is adopted, they wouldn't all have to essentially revalidate or update their indications for use with that triage strategy that would have already been approved. I don't know if that answers your question.

DR. VAN DER POL: So if I can restate that and make sure that I understand your answer, what you're saying is that the FDA or some entity would evaluate how well the triage process was working but that each new device would not necessarily be evaluated in that context because they're not going to ask for reevaluations of existing devices that got approval with older strategies, right? Okay. Can we have the risk-benefit slide up because I think that was a really helpful one as well?

DR. ZARITSKY: Sure.

DR. MOSCICKI: I just wanted to comment. Overall, I really think it's a great idea to do this because this is such a moving field. I always feel a little queasy when it just says professional, you know, organizations because this could be, you know, homeopathic or anything else. But all respect, Walter, even though everybody else was using every 5 years co-testing, Kaiser said we're going to do our own, we're going to do it every 3 years.

So, you know, I think even if we had specific indications, the way we practice medicine here in the United States, nobody follows it. So I think it really -- I think you're much better off, even from a risk thing, to say you need to follow professional societies. So I kind of went full circle, but I think overall I think it's a positive road forward.

DR. KINNEY: Walter Kinney.

If you could get agreement from the group that the new intended use direction was the right thing, then we could haggle about wording. But if the group thinks that they don't want to disconnect test evaluation from triage evaluation, then that's a separate problem.

DR. VAN DER POL: So, yeah, triage and follow-up. And this kind of comes back to where we were earlier talking about, you know, are we really just trying to see how well

this HPV test detects HPV versus how well this HPV test predicts long-term outcomes? So the intended use, then, this is still a related topic is what I was saying earlier, we would come back to this because this is -- and that's where we're coming back to it, so we --

(Off microphone comment.)

DR. VAN DER POL: Yeah, yeah.

DR. PERKINS: I think it is really important to evaluate HPV tests as screening tests because of the --

UNIDENTIFIED SPEAKER: For CIN3.

DR. PERKINS: For CIN3. Because of the -- because you can't have a test -- if it's available as a screening test, you can't -- or sorry, if it's available only as a triage test, then it's going to -- then you would require a lab to have two tests, one for triage of an abnormal and another one for screening, and that doesn't really make sense, and there's a lot of potential, I think, for misapplication if you have an HPV test that's only approved for triage.

DR. VAN DER POL: Maybe you can go back to that IFU because it wasn't only for triage. This is the proposed one, so if everyone wants to see that up there, it's a little hard to read. Can everyone read it, or should I read it out loud?

DR. PERKINS: Sorry. So I was actually agreeing with this, that I think having everything be approved for screening makes a lot of sense.

DR. ZARITSKY: And I do want to -- I'm sorry. I do want to point out that this is just an example of the wording and that if anybody has any modifications, you can discuss.

DR. FELDMAN: So in general -- this is Sarah Feldman.

In general, I like the idea of just approving tests for one purpose because that certainly would make it much simpler to then expand the use for other purposes; however, the underlying prevalence in the population and the prior probability of a test being positive is going to affect its usefulness or its sensitivity and specificity. So it is sort of

important to set the population correct that you're studying on so that you feel that it can be applied to a variety of circumstances.

DR. VAN DER POL: I think -- Nicolas.

DR. WENTZENSEN: I mean, it's getting back to my previous point, like the statement, women who test positive or negative for the high-risk HPV types, and then you're listing specific types, should have the following, and then later you have a triage test who supersedes that, I mean, how will you handle that on your end?

DR. ZARITSKY: So do you think that listing the specific types would be problematic in an indication for use like this?

DR. WENTZENSEN: It could result in having differential like recommendations, like here you would send like certain types to colposcopy, and the other setting you wouldn't do that and then like -- so it could lead to confusion, so I'm not sure how you could handle that.

DR. FELDBLYUM: So one possibility is we would have to list the subtypes just because that's what the device does, so we cannot not list it, but that's what the device is intended to do, to detect these specific subtypes. But then the language afterwards, let's say, should be instead of triage, we may not even need to use that word, we could just say that the patient should be followed up in accordance with professional guidelines and physician's assessment, and then it would be up to the physician to triage, not triage, follow up, whatever.

So it's just a matter of which language you want to use, but the idea is that we don't want to regulate clinical practice, we want to show that HPV test is one of the items in a complex algorithm, and the more complex the algorithm can get, the more different types of information are going to include, HPV is going to be just one piece of information. And so looking forward, that's what we're trying to do to see if we can approve an HPV test to

be one piece of information in whatever algorithm later is going to be defined.

DR. KINNEY: That seems eminently reasonable. Why wouldn't you want to do that? I mean, if there are objections, we should talk about them.

DR. VAN DER POL: I actually think Walter's point is valid, because if we all agree that the indication for use should be made more general, then it makes sense to wordsmith, and if we don't agree to that, then it doesn't make sense to go down that path, so maybe we could get nods, hands, thumbs up? We can do Girl Scouts.

(Laughter.)

DR. VAN DER POL: So do most people around the table think that it makes sense to at least look at how we would frame something like this? Can anybody give me some specific objections for why this is not a good idea? Not the specific words, but the concept as a whole.

Yeah, then Terri.

DR. DARRAGH: I'm into the wordsmithing part. I'll wait.

DR. DODD: Can I just -- this is Lori Dodd.

Can I just get some clarity on the subtyping and whether that would be included as part of the indication? Because I can see when you have a general population indication, who you actually evaluate the test on then matters a lot, right? If you tested in an unvaccinated cohort or a high-risk population and you have somebody who comes into your clinic who's extremely low risk and you don't know which type you're likely to get and you test a lower-risk HPV type and you're telling me I need to come in, you know, more frequently or something, it seems like as if one was a clinician and managing a patient, you'd want to know a little more detail if the indication is for general use.

DR. VAN DER POL: I think the way the -- this may be misleading because it says list types detected here, but that means what that test -- so like if you had a test that could pick

up types, and I'm just going to do this weird, if you had a test that could pick up Types 1, 2, and 3, because that's how you built your test, that's what would be listed there. But that's not to say Types 1, 2, and 3 follow -- so this is why it says follow the professional guidance because this package insert is not going to tell you what to do with that person, just whether or not they have that test, or that type.

DR. DODD: So I guess what I'm asking is are you saying that it can detect Types 1, 2, and 3, are you going to be reporting the diagnostic performance of each of those subtypes? Because what if the test has more affinity for Type 16 and -- but you don't know, when I come in with a Type 35, that that's what you're picking up, that's what I'm getting at.

DR. FELDBLYUM: So diagnostic performance will be evaluated either using the currently designed clinical study or a new clinical study approach that we discussed, but anything that goes in the intended use is evaluated, and then the performance estimates would be presented in the package insert or labeling. It's just not going to be in the intended use. But in the intended use, the device is designed, as Bobby said, to detect Types 1, 2, 3, then that's what would be listed, and then further down, in the package insert, you would have performance tables for those subtypes.

MR. SPRING: But just to -- this is Brad Spring, BD -- to clarify, I think there is a difference between detected and reported, and that may be the issue here, right? What gets reported to the clinician, is it all high risk in general, or are you reporting this is a 16, this is an 18, and so forth?

DR. VAN DER POL: Well, that's a different question, then, right?

MR. SPRING: Yeah. So we're talking about just detection here and not reporting, correct? Okay.

DR. DODD: I guess all I'm saying is that when you're going for a more general population, it seems like there's a little more responsibility to report more about the

population that the test has been studied under.

DR. MOSCICKI: So I just wanted to comment -- Barbara Moscicki -- that that is, I think, how we're trying to move forward in the NCI/ASCCP risk management that also involves -- so there are risk modifiers that will say what kind of population: screened/unscreened, vaccinated/not. We, in fact, have a whole committee looking for that risk modification. And so that would hopefully be those professional guidelines that look at risk; not necessarily all women are equal.

DR. DODD: Thank you. So would that information be in the indication for use or in the brochure?

(Off microphone response.)

DR. DODD: I guess what I'm -- in the brochure where you're reporting, presumably, those risk factors would be reported.

DR. MOSCICKI: Well, no. That's why I'm saying that it would be the professional guidelines. The professional guidelines will include those modifications, but they wanted to leave this more generic for indication, whereas that's where the professional guidelines come in.

DR. KINNEY: Right, this is --

DR. VAN DER POL: So this is --

DR. MOSCICKI: Which is a tricky wicket.

DR. VAN DER POL: So this may even say -- in your intended use, it may say I can pick up Types 1, 2, and 3, and the guidelines may say we only want to treat women who have Type 3, and if the way your test reports out as a pooled positive or negative result without types, maybe you'd choose not to buy that test.

DR. DODD: Okay. And I think I'm not speaking clearly, too. I appreciate that comment. What I'm saying is in the studies they conduct, surely you would want to report

the distribution of the risk factors that were captured in the study that -- in the sample population, right? I just want to make sure that -- that's all I'm saying. I mean, the risk factors are going to change, knowledge changes, but it seems like you would want that.

DR. MOSCICKI: If I understand, this is just indications for use, not the report that they would come to the FDA with.

DR. BURK: Robbie Burk, Einstein.

I'm in favor of a general statement of indications for use, but -- so basically you're saying, okay, we detected HPV 1, 2, and 3, medical guidelines say that you're now at a high risk, you're -- we're going to either follow up or -- but it's not going to be part -- it is not going to be part of the FDA's requirement that the indication -- the indication is going to be determined by the clinician or the specialized groups that deal with these diagnoses, but then we're coming to the point that basically we have a device that's detecting HPV at some level, and then the medical societies, the physicians make the decision how to follow them up. And it comes back to, you know, the question about why isn't analytical sensitivity the key outcome for a medical device that then is going to be used by doctors, whether it be Epstein-Barr virus or anything else, you know, I understand that, and we've been kind of going around in circles a little bit.

DR. VAN DER POL: And the only thing I will say about analytical sensitivity is for all of us -- of those who have done it and we work in matrix and we spike even whole buds, not just plasmids, they perform differently than patient specimens. And somebody said this morning, I think it was Dr. Schiffman, but somebody said this morning that, you know, there are all kinds of things in that specimen that we cannot always recreate when we buy matrix or we even try to collect matrix, so just doing analytical work and never collecting clinical specimens is going to be insufficient in the long run.

DR. MASSAD: Just for clarification, with a general indication for use, you're

determining that the device is safe and effective for detection of high-risk HPV. It's not that it's safe and effective for screening or safe and effective for triage or safe and effective for primary screening or co-testing or whatever.

DR. ZARITSKY: Yeah, so it would be -- I'm sorry, this is Luna Zaritsky from FDA.

Yeah, so it would be that this device is safe and effective as determined by whatever evaluation that we do, so it could be determined by having non-inferior clinical performance for detecting CIN3 or whatever the evaluation is, but yes, it would be just safe and effective for HPV rather than for triage or --

DR. MOSCICKI: But that's not what that says. It says cervical cancer screen to assess for risk of cervical cancer.

DR. ZARITSKY: Right, so the intended use will be supported by whatever evaluations are done, so this is just an example, and if the evaluation that is done looks at CIN3 or risk for CIN3 as an endpoint, non-inferior performance for CIN3+ as an endpoint, then that essentially would be assessing for risk.

DR. VAN DER POL: Did you have a question? You're shaking your head. Okay.

DR. WENTZENSEN: I have a follow-up question. Nicolas Wentzensen from NCI.

Previously, genotyping, like individual reporting of genotypes was tied to specific indications in the trials and in the evaluation. Now, I'm not sure if this -- like would this allow now somebody to come out with like a 12-type genotype test that can report out each type individually and throws it at the clinician and says, okay, now deal with it for the -- I'm like --

DR. ZARITSKY: Yes, so --

DR. WENTZENSEN: -- I would -- a little strongly but I think on one hand I'm excited, on the other hand, I'm worried about what that will do when -- and who's -- like how can we make sure this is done right?

DR. ZARITSKY: So if there is a manufacturer that comes out with a test that does offer, you know, all 12 or all 14 genotypes separately, then because that is going to be a paradigm shift in the screening field, we would definitely consult with experts in the field to make sure that this wouldn't have some sort of negative impact on screening, because if the overwhelming majority of the field are saying, no, this is a terrible idea, this would only jeopardize patients, then obviously we would take that into consideration, but we wouldn't make that decision in a vacuum.

DR. VAN DER POL: There was a question over here? No.

(No response.)

DR. VAN DER POL: So we've all agreed that we think that --

DR. BURK: So when you're evaluating a test for analytical sensitivity that you said you do, so if you use HPV 16, do you use only one variant lineage or do you use multiple variant lineages? And how does the genetic variations, most of these are DNA-based tests, how does genetic variation in the HPV take -- how do you take into account known variation and known variation in different populations?

DR. ZARITSKY: So I want to point out that when we do these analytical studies, that it's all done at the clinical cutoff, so we don't have an analytical LOD. But to answer your question is that we do not take into consideration all of the different isolates and all of the different genotype variations. For 16, for example, there's usually a plasmid that sponsors will use and a cell line. So we don't really take into account that level of detail for the genotypes.

DR. BURK: So that could have consequences in certain populations.

DR. ZARITSKY: Yes.

DR. VAN DER POL: So I want to go back and confirm, because I'm not 100 percent clear that we are in consensus, having a more generic IFU but still the language being based

on study design, so if the study were designed to follow up for CIN3, then it would be an IFU that had language like this where you're talking about cervical cancer screening versus if a study design were perhaps just HPV versus HPV, you would have an IFU that states that it detects HPV because that's what you studied, are people in favor of a more generic thing that pulls away the current descriptions of triaging and follow-up so that we can keep up with the fluidity?

DR. MOSCICKI: So I guess I'm a little confused because we just spent, I think, the last hour and a half discussing that CIN3 should be an endpoint. I didn't hear ever once anybody just say, oh, this HPV test is negative and positive, just like that one, and we don't want a clinical outcome. So I'm just a little confused now on why -- when would we ever not say for risk for CIN3 or cancer.

DR. GRAVITT: It seems to me that the question is can you do that with just almost like a case-control design of looking at test agreement among prevalently detected CIN3+, or do you need prospective evaluation of risk of future development as well?

DR. PERKINS: This is Rebecca Perkins.

I would argue that you do because you will never be able to control whether a test is used only for triage or surveillance, when does it become screening, so I think you have to assume that any HPV test has the potential to be used as a screening test, and so I think it should be evaluated as a screening test.

DR. GRAVITT: So Patti Gravitt again.

I guess I just conceptually don't know how, if you were to take a population which has a certain proportion where you're comparing two different HPV diagnostic tests, part of your population is just a general screening population, a few who will be detected with CIN2+ but probably not enough to power the study for that, and then you've got a separate population you're recruiting from colposcopy, which is going to be enriched for your

outcome, and you say, okay, I do a comparison between my investigative test and my comparator test within the screening population and I see reasonable concordance which doesn't suggest one's better than the other. To me, that indicates that you've got reasonable similar specificity. And you do a stratified analysis of the comparator and investigative HPV test and say, okay, that -- you know, the concordance between those tests within the CIN2, 3, however you want to define it by whatever biomarkers you're going to use to clarify your endpoint, they're similar, right?

* So to me, that -- I guess the question is, isn't that sufficient to say that this is clinically performing with a similar sensitivity and specificity as the already approved tests without having to prospectively follow, which we've already done with the comparator tests? Is it going to suggest that there's something uniquely different about the discordants that are -- so we're saying that the discordant samples in the general screening population would be randomly distributed if they're considered to be non-inferior. So to me, you only need to do prospective evaluation if you thought there was something very specific about the discordants in the investigative test that would have a unique risk relative to the comparator test, investigative test negative.

DR. FELDMAN: This is Sarah Feldman.

I think we're just getting confused right now. I think we're not talking about what type of study. We're just talking about the wording for this. And so the wording for this, all we're saying is that we would take out the specific plan that needs to be followed in order to confirm that diagnosis of CIN3, so that's really all we're saying, and so I think we sort of could use your study design or we could, in a different study, use a prospective.

But all we're really trying to say in this indications for use, this test can pick up for HPV, let's say, 16, which, you know, is a risk factor for cervical cancer, please follow professional guidelines for management, something like that. And I think that that is

valuable because some of the current HPV tests, you have to -- allegedly, you have to follow exactly the algorithm that they did in their research study in order to use it according to FDA guidelines, which is sort of -- doesn't make that much sense because we want to expand who has access without always having to follow exactly the research guideline.

DR. VAN DER POL: Right, thank you, because that is bringing us back on topic, and I think the study design question is one that's just thorny and will have to be another panel's headache, I think, sorry. But I think that we all do agree that an IFU that is not tied to current practice guidelines is probably wise.

(Off microphone comment.)

DR. VAN DER POL: Right, right. But just to say in there but not to specifically call out any current practice guidelines, because that is shifting sand, right, but to make sure that you say in there -- and practice guidelines and physicians' assessments in those instances where practice guidelines don't exist.

DR. FELDMAN: I just have one final question. Sarah Feldman.

So I asked this at the very beginning of the panel, but I'm going to ask it again. So suppose that something is approved now for clinical -- I mean, for clinician-collected swab or whatever and make an HPV positive test, is there anything about this that limits whether it can now be done -- I mean, self-administered in urine, you know, every month on your tampon or whatever the person feels like doing, is there anything about this wording that we could do to limit it from like random use?

DR. FELDBLYUM: Yes, there is definitely a limit in language, and if you look at the intended use of currently approved devices, it specifically states how the sample has to be collected and what collection media, and that will remain because a sample is a sample. It's -- apparently, I'm not pushing it hard enough. So as I said, the indication -- the definition of what type of sample and what type of media can be used will remain as it is in the

indications for use now. So if it says clinician collected using a brush or in ThinPrep or whatever media, it certainly does not mean that it can be used on urine or self-collection, so that limitation will remain.

DR. VAN DER POL: And I think the Panel agrees that that's wise. Can I have nods about that as well?

(Off microphone comment.)

DR. VAN DER POL: I don't think there's a lot of conversation about that either, but yeah. So it wouldn't be changing, really, any of the normal things that you see with how a test is used, right, in terms of specimen collection, specimen types, specimen media, so on and so forth, but it would just change how you manage the outcome, the results of that test, and to disentangle that from guidelines. Are there other thoughts about this?

(No response.)

DR. VAN DER POL: So I think this one's very clear. And I think that -- I think that we could go back and revisit study design, and actually, like I said, we could spend an entire day laying out study designs, right, but that's the purview -- sorry, Brad. That's the purview of the manufacturers, to get into that kind of thought process, and I don't think that that's really where we need to go here, but at least hopefully you've heard a fair amount of stuff about that. So, in general, I think we've had fair consensus on most of these topics, and I hope that's been helpful.

Does the FDA have any questions for the Panel?

DR. FELDBLYUM: No, we just wish to thank the Panel for their thoughts and active participation. This was very helpful. Thank you.

DR. VAN DER POL: I have to find my script, hold on.

DR. DARRAGH: Can I make some suggestions for wordsmithing before we go on?

DR. VAN DER POL: Certainly.

DR. DARRAGH: So Teresa Darragh.

I would change cervical dysplasia to cervical precancer and cancer. I would change professional guidelines to professional medical guidelines; there's lots of different professional societies. And I would change physician to something more generic like healthcare provider to allow for non-physician clinicians.

(Pause.)

DR. VAN DER POL: Okay, so I would like to thank the Panel, the FDA, and our guest speaker, Dr. Mark Schiffman, for their contribution to today's Panel meeting.

Do you have any final remarks on the part of the FDA?

DR. SCHERF: Yeah, this is Uwe Scherf.

Just a couple of words again. The discussion today demonstrated that it was the right time to bring these questions up, also that the questions were actually also right on target because we realize that this discussion today brought some controversial points up to light, and I think you provided us with clear comments and feedback that helped us to move to the next steps and do evaluation of the upcoming devices that are in our hands soon. This is very helpful. We see it from FDA perspective, this was a very successful panel meeting because you provided us with that information that is not easy to ascertain quickly, and I think really getting together here allowed us to do that, so thank you for that.

Again, thank you for your service, and have a very safe trip back home.

DR. VAN DER POL: Thank you.

So this meeting of the Microbiology Devices Panel of the Medical Devices Advisory Committee is now adjourned. If you would just leave your papers on the table, someone will collect those and manage them appropriately, so all of your handouts from your folder you can just leave right on the table. Thank you.

(Whereupon, at 3:42 p.m., the meeting was adjourned.)

C E R T I F I C A T E

This is to certify that the attached proceedings in the matter of:

MICROBIOLOGY DEVICES PANEL

March 8, 2019

Gaithersburg, Maryland

were held as herein appears, and that this is the original transcription thereof for the files of the Food and Drug Administration, Center for Devices and Radiological Health, Medical Devices Advisory Committee.

_____

TOM BOWMAN

Official Reporter

Free State Reporting, Inc.
1378 Cape St. Claire Road
Annapolis, MD 21409
(410) 974-0947