UNITED STATES OF AMERICA

DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOOD AND DRUG ADMINISTRATION

+ + +

CENTER FOR DEVICES AND RADIOLOGICAL HEALTH

MEDICAL DEVICES ADVISORY COMMITTEE

+ + +

HEMATOLOGY AND PATHOLOGY DEVICES PANEL

+ + +

October 22, 2009
8:00 a.m.

Hilton Washington D.C. North
620 Perry Parkway
Gaithersburg, Maryland

PANEL MEMBERS:

DOROTHY M. ADCOCK, M.D.	Chairperson
PIOTR KULESZA, M.D., Ph.D.	Voting Member
GEORGE G. BIRDSONG, M.D.	Temporary Voting Member
STEPHEN M. HEWITT, M.D., Ph.D.	Temporary Voting Member
JOHN H. SINARD, M.D., Ph.D.	Temporary Voting Member
JOHN R. GILBERTSON, M.D.	Temporary Voting Member
DAVID J. FORAN, Ph.D.	Temporary Voting Member
CLAUDIA MELLO-THOMS, Ph.D.	Temporary Voting Member
DIANE D. DAVEY, M.D.	Temporary Voting Member
XIAO-HUA ANDREW ZHOU, Ph.D.	Temporary Voting Member
DANIEL J. BRACCO	Industry Representative
LOUISE E. MAGRUDER	Executive Secretary

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

FDA REPRESENTATIVES:

ROBERT L. BECKER, JR. M.D., Ph.D.
Chief Medical Officer
Office of In Vitro Diagnostic Device
Evaluation and Safety

GERETTA WOOD
Director, Advisory Committee Staff

FDA SPEAKERS:

TREMEL A. FAISON, MS, RAC, SCT(ASCP)
ALDO BADANO, Ph.D.
ANANT AGRAWAL, MSE
MAX ROBINOWITZ, M.D.

GUEST SPEAKERS:

MICHAEL R. DESCOUR, Ph.D.
ULYSSES J. BALIS, M.D.
MICHAEL J. BECICH, M.D., Ph.D.

PUBLIC SPEAKERS:

GEORGE NETTO, M.D. for JUAN ROSAI, M.D.
MAHTAB FATEMI, J.D.
DIRK SOENKSEN
OLE EICHHORN
ROBERT DUNSTAN, DVM, MS, DACVP
GENE CARTWRIGHT, Ph.D.
LASZLO FONYAD, M.D.

INDEX

INDEX

INDEX

M E E T I N G

(8:00 a.m.)

DR. ADCOCK: Good morning. I would like to call this meeting of the Hematology and Pathology Devices Panel to order. I'm Dr. Dorothy Adcock, the Chairperson of the Panel. My area of expertise is coagulation. I'm the Medical Director of Esoterix Coagulation in Colorado.

Ms. Magruder, the Executive Secretary, for the Hematology and Pathology Devices Panel, will make some introductory remarks.

MS. MAGRUDER: Good morning, everyone. I will now read the conflict of interest statements for this meeting.

FDA Conflict of Interest Disclosure Statement (Particular Matter of General Applicability).

Hematology and Pathology Devices Panel of the Medical Devices Advisory Committee. Date of Meeting: October 22 and 23, 2009.

The Food and Drug Administration is convening today's meeting of the Hematology and Pathology 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 and Section 712 of the Federal Food, Drug and Cosmetic Act are being 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 who have financial conflicts when it is determined that the Agency's need for a particular individual's services outweighs his or her potential financial conflict of interest. Under Section 712 of the FD&C Act, Congress has authorized FDA to grant waivers to special Government employees and regular Government employees with potential financial conflicts when necessary to afford the Committee essential expertise.

Related to the discussions of today's meeting, members and consultants of this Panel who are special Government 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 purpose 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 on public health issues related to the use of digital whole slide imaging systems to replace conventional light microscopy for diagnostic surgical pathology. In the scope of this meeting, digital pathology is defined as capturing what can be observed in histologic glass slides into digital images via scanning devices; viewing these images via a computer monitor to render pathological diagnosis of the lesion of interest; and digitally archiving and retrieving these images. The meeting will focus on routine hematoxylin and eosin stained histologic glass slides. CDRH is seeking input from the Panel on several issues -- benefits and risks of the technology, study design, performance characteristics, and limitations of digital imaging hardware and software. Quantitative image analysis, computer aided diagnostics, and digital enhancements are not within the scope of this meeting. This is a particular matter of general applicability.

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 connection with this meeting. 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.

Mr. Dan Bracco is serving as the industry representative, acting on behalf of all related industry, and is employed by Constitutional Medical, Inc.

We would like to remind members and consultants that if the discussions involve any other products and firms not already on the agenda for which a 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 may have with any firms at issue.

Now, there are a few general announcements I would like to make.

If you haven't already done so, please sign the attendance sheets that are at the registration table by the door.

Transcripts of today's meeting will be available from Free State Court Reporting, and their number is (410) 974-0947.

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

I would like to remind everyone that members of the public and the press are not permitted around the Panel area, which is the area beyond the speaker's podium.

The press contact for today's meeting is Peper Long.

I would request that 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 AnnMarie Williams. AnnMarie, would you please stand.

Today's agenda will consist of presentations from our three invited guest speakers and five FDA personnel. The open public speaker session will immediately follow lunch. We have seven open public speaking requests for today. The rest of the day will be devoted to Panel discussions.

Unfortunately, due to unforeseen circumstances for which no time was allowed for us to find a replacement, we will not have a Consumer Representative at this meeting. Thank you.

DR. ADCOCK: Thank you, Louise.

At this meeting, the Panel will discuss and make recommendations on public health issues related to the use of digital whole slide imaging systems to replace conventional light microscopy for diagnostic surgical pathology. At issue is whether pathologists using digital whole slide imaging, given its current technological state of the art, is as safe and effective as examination and diagnosis of H&E stained glass slides by pathologists using conventional light microscopy to render pathologic diagnoses of routine surgical specimens. FDA also would like input on the necessary elements and methodology of a sound clinical investigational plan to validate that digital imaging of H&E glass slides by a whole slide imaging device.

Before we begin, I would like to ask that our Panel members and FDA staff seated at the table, to introduce themselves. As you do, please state your name, your area of expertise, your position, and affiliation. We'd like to start with Mr. Bracco.

MR. BRACCO: Good morning. My name is Dan Bracco. I am the Industry Representative, and I'm currently Vice President of Clinical, Regulatory and Quality with Constitution Medical, Incorporated. Thank you.

DR. DAVEY: Hi, I'm Diane Davey, and I'm a pathologist. I'm a Professor of Pathology and Assistant Dean at the University of Central Florida, and I also work part-time at the Orlando VA Medical Center, and my areas of expertise are cytopathology and hematopathology. Thank you.

DR. BIRDSONG: I'm George Birdsong. I'm a pathologist at Emery University School of Medicine in Atlanta, and I'm Director of Anatomic Pathology at Grady Hospital. My area of expertise is cytopathology and informatics as related to anatomic pathology.

DR. ZHOU: Xiao-Hua Andrew Zhou. I'm a biostatistician. I'm a Professor of Biostatistics at the University of Washington, Seattle, and also the Director of Biostatistics at the Seattle VA Medical Center.

DR. HEWITT: Lieutenant Commander Stephen Hewitt, United States Public Health Service, appointed as a clinical investigator at the National Cancer Institute where I serve as Chief of the Tissue Array Research Program and Applied Molecular Pathology Laboratory. Special expertise is

high throughput pathology.

DR. SINARD: Hi, I'm John Sinard. I'm a Professor of Pathology at Yale University School of Medicine. I'm the Director of the Informatics Program, Director of the Autopsy Service. I do general surgical pathology as well as autopsy pathology.

DR. GILBERTSON: John Gilbertson, Associate Chief of Pathology at Mass General, Associate Professor at Harvard University, and Director of the Informatics and Imaging Fellowship Program at the Department of Pathology at Mass General.

DR. MELLO-THOMS: Hello, I'm Claudia Mello-Thoms, and I'm an Assistant Professor of Biomedical Informatics and Radiology at the University of Pittsburgh. My area of expertise is image perception.

DR. KULESZA: Piotr Kulesza. I'm an Assistant Professor of Pathology at Northwestern University in Chicago. My expertise is cytopathology, and I also run a research laboratory in experimental therapeutics and oncology.

DR. FORAN: I'm David Foran. I'm a Professor of Pathology and Radiology at Robert Wood Johnson Medical School and Director of Biomedical Imaging at the Cancer Institute of New Jersey. My areas of expertise are pattern recognition and image analysis.

DR. BECKER: I'm Bob Becker. I'm an APCP Board Pathologist. In a former life, I was involved in research and development in digital imaging

and pathology. I'm currently the Chief Medical Officer for the Office of *In Vitro* Diagnostic Device Evaluation and Safety at FDA.

DR. ADCOCK: Thank you. Ms. Geretta Wood, Director, Advisory Committee Staff, would like to say a few words.

MS. WOOD: Good morning. At this time, I'd like Ms. Magruder to please join me at the podium. Today, I would like to recognize Ms. Louise Magruder. Louise has been at the FDA since 1990. She has been an Executive Secretary of this Panel since 2002. In addition, she also served as the Executive Secretary for the Immunology Panel.

I am also happy to say that Louise has been a colleague and friend of mine since I came to FDA. Because of the reorganization of the Advisory Committee Staff and the centralization of the Executive Secretary responsibilities and also because Louise is first and foremost an extremely knowledgeable product premarket reviewer, she has decided to devote her efforts within the FDA as a scientific reviewer.

So today, I would like to show my appreciation and the appreciation of the FDA by presenting Louise with this plaque and by offering her my extreme gratitude for her dedication and service as the Executive Secretary serving in the Advisory Committee Staff. Thank you, Louise.

MS. MAGRUDER: Thank you very much.

(Applause.)

MS. MAGRUDER: I would just like to say that it's been

extremely rewarding serving as Executive Secretary for the Advisory Panels. It's a privilege to meet and to work with this particular Panel, and I thank you for that honor. I also just want to tell you all how important the Panels are to the FDA regulatory process. Thank you.

MS. WOOD: I would also like to take this opportunity to introduce the incoming Executive Secretary for the Hematology and Pathology Panel. Ms. Margaret Janicki graduated from Mary Washington College in 2000. She holds a Bachelor of Science degree in Chemistry with a concentration in German. Margaret has worked at the FDA Center for Biologics Evaluation and Research for six years, and she has served as CBER's Institutional Animal Care and Use Committee Coordinator for the past five years.

Margaret, please stand.

(Applause.)

MS. WOOD: We welcome you, Margaret, and we all look forward to working with you. Thank you.

DR. ADCOCK: Thank you. The first FDA presenter is Ms. Tremel Faison, Scientific Reviewer in the Office of In Vitro Diagnostic Device Evaluation and Safety, Division of Immunology and Hematology Devices. She will give the historical overview of FDA regulation of digital pathology imaging applications. Thank you, Tremel.

MS. FAISON:

Good morning. My name is Tremel Faison, and I'm a Scientific Reviewer in the Office of In Vitro Diagnostic Device Evaluation and Safety. I'd like to welcome the Panel to this meeting regarding digital pathology. I, as I'm sure you are, am looking forward to a productive and informative two days.

Advances in technology, such as improved image software, scanning, throughput, and image storage capabilities have made the widespread use of digital imaging systems for primary diagnosis a reality; if not now, in the near future.

Today, I'd like to give you a historical overview of FDA regulation of digital imaging devices in OIVD. The use of digital imaging is not new in the clinical anatomic pathology laboratory. OIVD has cleared and approved several digital imaging devices with limited adjunctive applications.

We believe the use of digital imaging for surgical pathology raises new issues safety and effectiveness that must be addressed. We recognize the many benefits the technology provides. At the same time, we need to be sure of its limitations to present risk to public health.

So I'll define digital pathology and the intended use for the scope of the meeting. Next I'll talk about the historical reference standard, the light microscope. I'll discuss some examples of IVDs that utilize digital imaging and their intended uses.

And to switch from digital film to digital radiology, while not

totally applicable to digital pathology, can provide some valuable lessons for us to draw from.

Last, I'll discuss our concerns about digital pathology devices and the associated safety and effectiveness issues.

Digital pathology is defined as the use of computer technology to convert analog microscopic images into digital imagining. Whole slide imaging, or WSI as we will refer to it, is also known as digital imaging, virtual slides, virtual microscopy, and I'm sure we can come up with a lot of other aliases.

Whole slide imaging systems consist of hardware, a microscope, camera, scanner, computer and monitor, and the associated software. The system encompasses image acquisition, processing, archiving, and retrieval.

So everything here at the FDA starts with an intended use. Intended use shapes how the device will be regulated, in what population and ultimately the study design.

For the scope of this Panel meeting, the intended use is defined as the use of whole slide imaging for primary diagnosis of surgical pathology microscope slides in lieu of a microscope. This is not an adjunctive intended use, and for our purposes, we will be considering the use of whole slide imaging for all surgical pathology specimens. It will not be organ or disease specific.

The historical reference standard for surgical pathology

diagnosis is the light microscope. They are defined as medical devices under the Code of Federal Regulations or C.F.R., under 864.3600, as optical instruments used to enlarge images of specimens, preparations, and cultures for medical purposes.

Microscopes are Class I, exempt from premarket notification subject to the limitations in 864.9.

So 864.9, limitations of exemptions from 510(k), and here are just a few. Exemption is only to the extent that misdiagnosis as a result of using the device would not be associated with high morbidity or mortality. If there's a different fundamental scientific technology or as an IVD intended for use in diagnosis, monitoring, or screening of neoplastic diseases, the device is not qualified for exemption.

So what does all this regulatory language mean? Here is a graphic of each of the components of a whole slide imaging system. As you can see, the optical microscope is only one part of the system. We consider other components such as image acquisition, processing and display new technology for the intended use.

In addition, the use of whole slide imaging for primary diagnosis of surgical pathology specimens encompasses not only neoplastic disease but the entire spectrum of human pathologic diagnoses and carries the risk of serious public health consequences if erroneous results are rendered using these systems.

Therefore, whole slide imaging cannot be considered exempt from premarket submission requirements.

So if whole slide imaging systems aren't Class I, then are there any other digital imaging devices in IVD that we've had past experience with?

Automated cell-locating devices are regulated as Class II and encompass a variety of different types of devices. Some examples here include automated hematology analyzers with differential cell counters that count cells and blood or body fluids, chromosome analyzers seen here in the upper right, and FISH enumeration systems in the lower right, and also urine sediment analyzers.

What all of these devices have in common is that they locate, count, and identify or classify cells, and they must have a final review or interpretation by a trained user.

We also have image analysis devices. They've been used as an adjunct to light microscopy diagnosis to quantify the amount of brown stain for immunohistochemistry stains, such as for HER2/neu and estrogen. These devices are only used after the primary diagnosis is made under the microscope and are used by a trained pathologist to review the results.

We also have digital images that are used for immunohistochemistry with no image analysis. We call this a manual read of a digital image, and the pathologist performs an estimation of the stain intensity and percent positivity from the digital image alone. It is limited to

breast PR and HER2 applications.

Last, we have gynecologic cytology imaging systems. These systems use detection algorithms and are neural networks to detect abnormal cells on pap smears. They are Class III devices that aid in the identification of abnormal cells by cytotechnologists but perform no diagnosis. In fact, diagnosis is still made on the abnormal cases under the light microscope.

So none of these examples can be directly applied to digital whole slide imaging. Each has a specific and limited intended use that is not applicable to whole slide imaging for the breadth of surgical pathology specimens.

So we concluded that whole slide imaging raises new questions of safety and effectiveness that must be answered through premarket submission requirements.

So what are the FDA's concerns? We're concerned that if the whole slide image is presented, is it of such quality that the same diagnosis could be made as when using the light microscope for all surgical pathology specimens?

What are the differences in use between the two methods? What is the difference between viewing and navigating on a computer screen versus under a microscope?

There could be serious consequences to public health if misdiagnosis is caused by poor quality image or improper use.

So what can we learn from radiology? There are similarities between the switch from film to digital mammography that pertain to digital imaging. And there are several radiology regulations that can serve as examples. I won't go into detail here. Tomorrow, Dr. Myers will discuss the safety and effectiveness issues and regulatory requirements for digital mammography. Regrettably, many of the tough issues associated with whole slide imaging are different and have more subtleties than digital radiology.

However, digital radiology does give us a nexus to begin discussion on how to determine the safety and effectiveness of whole slide imaging.

So how does FDA determine safety and effectiveness? We must consider the intended use population that the device will be used on, the conditions of use for the device, the probable benefit to health from the use of the device weighed against any probable injury or illness, and the reliability of the device.

Safety is defined as a reasonable assurance that, based on valid scientific evidence, that the probable benefits to health from use of the device outweigh any probable risks.

And effectiveness is a reasonable assurance, based on valid scientific evidence, that it will provide clinically significant results.

How does FDA plan to ensure the safety and effectiveness of digital pathology devices? We'll require analytical and clinical studies to

objectively and precisely validate performance. We need to know the risks, benefits, and limitations. We would like to see standardization of digital pathology, and if enough information isn't available from premarket studies, we may require postmarket studies. We're seeking input from the Panel on these issues.

In summary, we recognize that the technological advances associated with whole slide imaging make its use a reality. Whole slide imaging systems are not Class I exempt and are therefore subject to premarket requirements.

Current IVDs that utilize digital imaging for limited applications are not applicable to the whole slide imaging paradigm.

Digital mammography may provide useful lessons but does not address all of the concerns for whole slide imaging.

And our goal is to gain information about the technology in order to ensure safe and effective use.

I also have a note for the Panel members, that the questions that you will be presented with today have been condensed from the form that they were in, in the Panel packet.

DR. ADCOCK: Thank you, Tremel.

I'll introduce our next speaker. Our first guest speaker is Dr. Michael Descour, who is an Associate Professor at the College of Optical Sciences, University of Arizona. Dr. Descour will be speaking on light

microscopy, the physical characteristics and measurement of optical imaging.

DR. DESCOUR: Well, good morning, everyone. Thank you for the invitation to address the Panel. My purpose today is to introduce you to how a microscope functions, what makes it work given our current thinking on the subject, and then based on this explanation, how it can be evaluated in terms of its imaging performance.

The talk is divided into four sections. We'll begin first with an introduction, which I guess could colloquially be called the dark side of the microscope. We'll then talk about departures from ideal behavior known as aberrations in light microscopy. We'll then follow that with a brief description of several techniques used for evaluating objectively microscope performance, and then finally we'll conclude with a few remarks about how an image produced by a microscope appears to a human observer and perhaps conversely how a human observer might appear to the microscope in reverse.

So let's begin with a quick description of the anatomy of a modern microscope. There are several subsystems, most of which are probably very familiar to all of you. We'll begin first with the light source and associated optics that form the illumination system in a microscope. Then, of course, its whole purpose is built around the glass slide and more importantly for our purposes, a cover glass or a piece of tape. Then we encounter a microscope objective and finally an eyepiece, and then, of course, the human

eye to form the final image and actually observe it.

We'll discuss how images are formed in a light microscope based on a treatment way of thinking about this topic that dates back to the middle 1800s and Ernst Abbe, and then we'll conclude with a brief comment on lateral and axial resolution of microscopes again under ideal conditions and say a few words about how they might be degraded if there are departures from that ideal behavior.

So the story begins with a compound light microscope definition. That instrument essentially behaves like a very sophisticated magnifier. Its purpose is to present to the observer a magnified image which is projected to a comfortable viewing distance. The essential modification that the microscope accomplishes is that the human eye can comfortably focus on features and objects about 10 inches in front of it, and any closer than that, the eye is unable to focus on the object and any small detail that you might wish to observe in it. So the light microscope basically takes the small object and puts it in a magnified form at a distance where the eye can very comfortably focus on it and allow us to see very fine details within the specimen.

You can think about what the microscope does in sort of a reverse way, and that's what the chart on the right is intended to show. I guess I better get a more powerful laser pointer here. So just quickly to explain the structure of the compound light microscope, it's called that by the

way because it does the magnification in two stages. The first one is accomplished by a microscope objective represented here with just a single lens. The specimen is on the left. That microscope objective forms a real image, and we'll explain why we say real in just a short while, a magnified real image over here, which is then further magnified and projected to a comfortable viewing distance by an eyepiece. So that's the structure of the compound microscope.

You can think of what it does from a kind of different perspective, and that is that the microscope essentially takes the human eye and maps it in a way to make it capable of seeing very fine details very close up, and so the pupil of the human eye is projected to the aperture stop of microscope objective by the optics of the light microscope, and similarly the retina is projected to where the real image is formed by the light microscope. So in effect the light microscope maps the human eye into sort of an enlarged, more capable version of itself, capable of seeing very fine details. So that's the imaging part of the light microscope.

It turns out that there's also some sophisticated light control and collection going on, on the illumination side. Most modern microscopes employ a form of illumination called Kohler illumination. There are a couple of other types of illumination that are far less commonly used, critical or diffused. In Kohler illumination, the essential idea is that we use some kind of a light source, which could be in the form of a filament lamp, which will have

a lot of structure and potential texture to it, maybe -- uniform. So we prefer not to project the light source directly onto the specimen because in that case its structure would compete with the structure of whatever we wish to see in the specimen.

The way one gets around that is by Kohler illumination, and effectively what happens, let me just kind of walk you step-by-step through the steps of this process in the upper diagram. So we begin with the lamp, again, a variety of choices there including light illuminating diodes nowadays, but we'll stick to the filament idea. That gets imaged by a collector lens to the aperture called the aperture diaphragm, sometimes also known as the iris diaphragm. That can be opened or closed, and the adjustment that you make to this aperture will affect the type of illumination, specifically the numerical aperture of the illumination on the specimen. That can have a very profound effect on resolution, contrast, and depth of field, and we'll spend a few minutes talking about that adjustment a little bit later in the talk.

The specimen then, I'll just highlight it here, we'll come back to it in a moment, is located right there, that vertical blue line, but if we follow the image of the filament, next time we see it is in the real focal plane or the aperture stop plane of the microscope objective, and then finally it is imaged into the pupil of the eye. So that the main point here with the illumination part of this system is that an image of the filament is never very close to an image of the specimen. We put it about optically as far away as we can from

the specimen to make sure that the illumination at the plane of the specimen will be as uniform as possible or the image of the filament will be as defocused as possible of the specimen so that we get uniform illumination both in terms of irradiance or intensity if you prefer, as well as numerical aperture.

Now, the next path to follow here involves actually imaging the specimen. Now, in this case we encounter another diaphragm in the microscope, in the modern microscope, known as the field diaphragm. Its purpose is to limit the illuminated area on the specimen itself. Closing down this diaphragm basically shrinks the illuminated field of view on the specimen. So that diaphragm is illuminated, is projected or imaged onto the specimen plane, and then the specimen, of course, is imaged to an intermediate real image, sometimes called the primary image, and that is the image that is then going to be viewed by eye through an eyepiece and, of course, its final rendition appears at the back of the eye, at the retina.

So you can trace basically two sets of rays through modern illumination and microscope combination, and this form is pretty much the current optimized version of a light microscope.

Now, if you wanted to take these forms and turn them into a whole slide imaging system, essentially what you would do is remove the eyepiece and the eye, of course, and replace and put at this intermediate image plane a digital sensor of some kind, a CCD or CMOS or TDI sensor.

So many of the optics are shared between the two configurations. The differences really appear at the back end, what does the actual imaging.

Okay. So now we get into a little bit of theory on how images are formed in a microscope. I want to introduce this with the primary intention of sort of setting the foundation for how we might later evaluate the microscope, where we might look for any kind of departures from ideal behavior in this kind of an instrument, and what it might affect.

The theory actually dates back to the middle of the 19th Century, the 1850s, and it involves Ernst Abbe who was one of the prime scientists working for Karl Zeiss at the time, and his premise essentially was that first of all, we can think of a specimen, and here you will just have to take my word for it, that we can think of a specimen as a superposition of gratings of many different frequencies, at different orientations, and with different amplitudes. By amplitude, I mean how much light they transmit in the clear spaces.

Now, that treatment allows us to then, because it's a linear superposition, we can think of how a microscope would treat an individual grating, some of those effects, and get an effect of how the microscope would treat a specimen of some general nature.

So the illustration below attempts to explain what would happen if we were to look at just a single one of those gratings that represent,

that can be added up to form the object, in a regular microscope, and you can actually observe this yourself if you remove the eyepiece and peer down at the aperture stop of the microscope if you happen to have a grating handy.

So the object grating is essentially just a bunch of lines, opaque and clear, that are illuminated with an illumination system. In this case, I've kind of idealized the illumination system to just one direction of illumination. The grating forms a series of diffraction orders, and that's what these up and down trending bundles are meant to represent, and then with the aperture stop, what we observe are three diffraction spots, each associated with one of these diffraction orders.

Now, if you want to know where this comes from mathematically, you can get it by doing a two-dimensional Fourier transform of the grating. We won't get into that at the moment, but the most important thing I want you to appreciate is the relationship between the frequency of this grating and its orientation and the displacement of the spots.

Essentially what happens if I increase the frequency of the grating, meaning I make the period of it shorter, these spots, the top one and the bottom one will move further and further away, and at some point, if the frequency is high enough, those spots will fall outside the limits of the aperture, and I will no longer be able to perceive the grating in the image plane of the microscope. But if the spots do fit within the extent of the aperture stop or the microscope objective, then what happens at this

intermediate image plane according to Abbe is that we have multiple plane ways and propagating at the image plane, and where they all interfere with one another, we then see this kind of a zigzagging profile or another words, an image of the grating produced via this kind of a process. And then, of course, after that we look with the human eye and the eyepiece at this grating and see that there was this particular object located in the specimen plane.

Now, where the imperfections of a microscope appear has to do with what happens in this dark circle here or within the aperture stop. If these diffraction orders encounter any kind of departures from ideal behavior in the microscope, that will then directly affect the quality or, in other words, the contrast and resolution of this grating image. So this description here forms essentially a foundation about how we think about microscopes and also how we can evaluate them, and I'll explain that a little bit later in the talk.

I should take a very quick detour here and go back to this aperture diaphragm that we talked about a moment ago. It turns out that in most microscopes, the size of the sources image, think of it as the filament or perhaps some kind of another type of light source, when you look down into the microscope by eye, and there's no object present in the specimen plane, you will see the image of the source indicated here of this white circle inside the aperture stop of the microscope objective.

Now, it turns out that in most microscopes, the source underfills the full dimension of the aperture stop of the microscope objective. The ratio in linear terms and radius terms tends to range anywhere from two-thirds to three-quarters, but this is a typical practical situation. Now, of course, you can open up the diaphragm and fill the aperture stop, or you can close it way down, and if you've ever experimented with that, you will see all sorts of diffraction patterns around the features on the specimen, but the point here I think that I want to convey to you, and this is really the bottom line message right there, that any evaluation of a microscope's resolution, contrast, or depth of field needs to be connected to whatever was the setting of this aperture diaphragm because as we change this, the fill of the aperture stop, we will affect each of these parameters in a way that may enhance some of them but perhaps degrade others, resolution and contrast being a perfect example of that kind of a tradeoff, of increasing one at the expense of the other.

So then let's say a few quick words about resolution of a light microscope because ultimately that's what it's all about with these instruments. I should first point out that resolution is sometimes confused with image pixel format in digital images. What I mean by resolution here is what level of detail is a microscope capable of seeing and telling me, resolving if you will, that there are multiple features in my specimen that may be very small and very close to one another.

The most common criterion, and by the way, these criteria area slightly arbitrary, but I'll just mention, the most common one, and then a couple of alternatives, the most common one is due to Lord Rayleigh. It's a rather old criterion as these things go. It dates back to 1896. It's essentially based on astronomy-based thinking. Here you are faced with a couple of point objects of equal intensity. So think of maybe two stars very close to one another, and the question becomes how close could those two stars be to one another for me to be able to tell that there are two of them as opposed to just one big blob.

The Rayleigh criterion recognizes the fact that a microscope will not render a point into a point but rather into a finite blur, and that's what you see with these two curves that I have superimposed on top of one another in the upper right, and when these two curves are added, we get a curve indicated in blue which has a dip in the middle and goes back up to another peak and then falls off again.

So this is a profile of the intensity and an image that would allow me in the limit to say that I'm actually looking at two separate sources as opposed to just one large diameter feature in an image.

Now, it turns out that this resolution criterion is tied directly to the wavelength at which the microscope may be operating, and this might be, for example, the longest wavelength at which we use a microscope, somewhere in the red, for instance, and the numerical aperture of the

microscope objective. So with these two parameters, and nothing else, you can determine what is this minimum separation that will be resolvable according to this criterion.

Now, it also indicates for us what is the sampling distance necessary. So the larger the minimum sampling distance that we can get away with, and it's essentially this distance D, this separation from this to this peak, divided by a factor of 2, and the reason for the factor of 2 is that we want to have a sample, and the sample here might mean a pixel located at one of these peaks, then one in the trough to tell us that there is a depression in the middle, and a third one at the second peak. So from this you can get an idea, given the wavelength and the numerical aperture value for a given operating scenario, what should be the size of the pixels in order to resolve these features in the limit according to Rayleigh.

Now, I mentioned there were some alternatives. One of them is so-called Sparrow criterion. Essentially, numerically speaking, what they all do is replace this factor of .61 with a slightly different number. So, for Sparrow, we replace it with .47, and the difference here is that instead of having a dip in the center, we are looking for flat tops, so we can put the two features a little bit closer together and get just a flat top in the combined intensity curve, and then Ernst Abbe also had his own criterion, and in that case, the number goes to .5. So it's a little bit of a self science talking about resolution, but all of these criteria basically strike a similar note, and they

each dictate something about the sampling size required to observe these objects in the limit of resolution.

The other thing I'd like to point out is that these criteria sort of indirectly apply to a light microscope in bright field imaging again because we're really talking about bright points and a dark background which is sort of the opposite of what most pathologists would be used to seeing, and also noise is not being considered in these criteria.

Okay. So far everything I've told you assumes that things behave ideally and there are no aberrations in the optical system, were it only so, but that would, of course, put all the lens designers out of business. So it turns out that there are these so-called aberrations. They are effectively departures from the gold standard of imaging, which is known as diffraction limited imaging performance, and diffraction limited means that we're limited essentially by fundamental properties of focusing light, and that has to do again with wavelength and numerical aperture.

There are two primary kinds of aberrations. I always get choked up when I talk about aberrations. So we have two primary kinds. The first are independent of position in the field of view, and a great example that I'm sure most of you are familiar with is spherical aberration that you might, for example, see if a slide is covered with the wrong thickness of cover glass. That aberration does not vary across the field of view indicated here with this larger circle, and these little rezones or these rings are intended to indicate

sort of the point spread function, what kind of a blur you would get from a point object. So the essential message here is that the blur is the same no matter where I go in the field of view.

The other kind of aberration is that which varies with position, within the field of view, and that looks something like this. In the very center of the field of view, the aberration is perhaps either a zero or a very limited effect, but as we move away from the center, the aberration gets worse. The effect of these aberrations is to limit the field of view size. So the top ones just blur the image everywhere the same. The bottom, the field dependent aberrations will cause a reduction of the field of view, the useful, practical field of view.

Now, things get actually more interesting after that because we also have aberrations related to the wavelength of light that might be passing through an optical system. There are two flavors of those. We'll talk about them in the next view graph, and then finally there is another kind of aberration which has a field dependence. It's called distortion, and in this case, it's a very particular technical meaning. It essentially has to do with the idea that magnification changes as we go from the center to the edge of the field of view. A very well known type of distortion is fisheye distortion with a fisheye lens. Typically these two are known as barrel distortion or pin cushion distortion, but you can encounter others.

That might be particularly important if you're actually trying to

derive precise measurements from an image, the distortion. Depending on where an object is located in the field of view, it might appear slightly larger or slightly smaller.

Okay. So let's just delve into this a little bit more. In the case of the aberrations that are independent of position within the field of view, a couple of obvious ones are defocused that affects all parts of the field of view the same, and spherical aberration, we talked about that already being tied, for example, to the cover slip thickness. I'll have a chart to show on that in just a moment. Axial chromatic aberration basically is a wavelength dependent or color dependent defocus. Different wavelengths will focus at a different position, and all of these aberrations collectively will basically blur the image in some uniform fashion across the entire field of view.

Again, in the case of the field dependent aberrations which depend on position within the field of view, we have things like field curvature. Objects in the middle may be in focus, but objects at the periphery of the field of view will be blurred out. Lateral chromatic aberration, which is a change of magnification with color, that can lead to color misalignment at the edges of the field of view -- astigmatism are also occasionally encountered. All of these aberrations will affect and reduce typically the useful size of the field of view.

Now, one of the things that we'll talk about in just a moment will be something called the modulation transfer function, but I'll give you one

hint of what's coming, and that is that when evaluating the imaging performance of a microscope, if you only measure to one position, you are effectively assuming that all the aberrations are independent of the field of view, which is not an accurate picture, and to really get a complete view of how well the microscope works, you need to repeat these measurements on multiple occasions within the field of view in order to understand its overall performance. And, of course, we talked about distortion already.

Now, what can we do to correct aberrations? Most modern microscopes do a very good job of handling these problems. Typically their correction is divided between the microscope objective itself and some other optical element in the microscope, and in the old days, it used to be the eyepiece. Nowadays it would be the tube lens, and so these two separate optical systems acts to correct the aberrations of one or the other, and usually it's the tube lens, for example, compensating lateral color in the microscope objective.

The message is that this one is careful and uses an objective all by itself, it may not actually be a fully corrected set of optics. It may not be a fully corrected optical system, and it will exhibit some residual aberrations that were expected to have been fixed by either a tube lens or an eyepiece which you may not have in an instrument. The result of this is that the aberrations of the objective then are effectively baked into a digital image, and you can actually find them in digital images if the system is being used

partially as described here.

Let me quickly say a couple of words about spherical aberration. By the way, throughout the talk, I've tried to pepper it with ISO logos to indicate where there are existing standards by the organization that have to do with some particular aspect of either microscopic imaging or sort of neighboring imaging type that might be relevant to our discussions here.

In the case of spherical aberration, that's one example where especially a high numerical aperture, SA by the way stands for spherical aberration, that effect becomes more pronounced as the thickness of the cover glass or cover tape departs from some design condition. So that's another important parameter for high numerical aperture objectives where the NA exceeds .5. So it might be like a 20X or 40X or higher magnification objective. Those optics typically are designed with a particular cover glass thickness in mind. The cover glass in effect is part of the overall optical system. If you replace the cover glass with one of a different thickness or a different index of refraction, it's as though you had cut open the objective and replaced one of its lenses with another lens, and I think as you can imaging, that may degrade the performance of the overall system. The same is true if we replaced the microscope cover slip with one of a different thickness. That amounts to use of the system off design, and it will typically degrade the imaging performance.

So there is actually a standard for cover glasses that define

some of these parameters. Microscopes today typically of high numerical aperture have some kind of a manual compensation adjustment where the user can turn a dial on the microscope barrel itself in order to correct for a different thickness of the cover slip for which the objective was originally designed.

Okay. So let's say a few words about how we could test microscope performance. All of these techniques are very basic, and the ones we're going to talk about again tend to have some kind of a corresponding ISO standard associated with it. So there's a well-described method for testing of optical systems that would also apply in the case of microscopes.

The first technique is a so-called tilted or slanted edge method. This is a very nice method because with just one image, we actually collect a whole bunch of information about the performance of a microscope. It's lateral resolution. We can also talk about how well chromatic aberrations are corrected in the system and also detect any kind of color misalignments.

For distortion, again this is a field dependent aberration, changes of magnification. That, too, can be tested with a special purpose pattern in place of a specimen. Same is true in the case of a field curvature. We'll say a few quick words about this, and color misalignment and images, again there are simple tests that can be executed using a special purpose specimen, sort of an artificial specimen, and then I'll just mention a couple of things about imaging continuity that might be specific to instruments that

stitch or put together a larger image out of smaller fields of view.

The one thing that's going to be noted immediately is that all of these tests can be performed using some kind of a special purpose target. Some of these targets are available commercially today. Others may need to be created, but it's really a rather straightforward process to make these kind of special purpose resolution targets. Again, you can think of them as artificial specimens.

Now, one other final point is that all of these techniques that we're talking about here are methods that involve taking a microscope, putting a special purpose object in place of a specimen, then either looking or recording the images that are forms of that specimen and evaluating them.

There's another class of tests based on a method called interferometry. We're not going to spend much time talking about these. They can be used, but these tests tend to be a lot more sophisticated and more challenging to carry out. So the methods that I've pointed out with the bullet list here would be techniques that could be executed very readily in the field as part of an evaluation or a calibration procedure.

Okay. Now, to introduce you to this slanted edge method that is probably one of the most informative techniques of testing a microscope, let me just say a few more words about this concept of a modulation transfer function. This is a very well established concept in optical engineering. It essentially can be thought of as evaluating how well a microscope or any

optical system in general can represent some kind of a grating, high contrast pattern located in the specimen plate. And it turns out that as the frequency of the grating changes, i.e., increases, the capability of the optical system to show you a high contrast image of that same grading begins to diminish, and in the limit, ultimately if you put in too high a frequency grating in the plane of the specimen, all you will get of the image plane will be uniform illumination. The instrument is no longer capable of representing any kind of contrast in that sort of an image.

Now, there is a certain ideal shape to this curve that would tell us how the optics can pass different special frequencies better than others, and then you have shapes of this curve that reflect the presence of aberrations. So this is a very simple and illustrative test of seeing how close an optical system comes to ideal performance limited again by diffraction as we said earlier.

Okay. So let me just walk you through this step by step. I hope everyone can see this. There's a bright tilted rectangle surrounded by an opaque region. So we're looking at a small section of an image of a particular artificial specimen that has a clear section, and that's kind of square shaped and then is otherwise coated with metal in this case to make it opaque.

So the first step is to collect this kind of an image, and the target has been tilted slightly for the following reason, that if we are taking this image with some kind of a pixilated sensor, it's going to have usually a

squared root of pixels on it. If the edge were not tilted, then we would be getting slightly different results depending on the orientation of the edge with respect to the pixel grid itself. By tilting the edge, we effectively remove that variable completely from consideration.

So as a result, what we get is something called the edge spread function, which is indicated by this experimental data over here, and this actually shows us what is the transition in this particular system from bright to dark. Now, you might have expected a very nice straight edge, but unfortunately because optics work as they do, we're going to get some degree of blurring, and this is known as the edge spread function. So whenever you hear about any kind of a spread function, that again reflects an optical system will spread out or blur some feature by some amount, either limited by diffraction in the ideal case or limited by aberrations in a non-ideal case.

So from this edge spread function, we can rapidly obtain just by taking the derivative of something called the line spread function, same concept, except now instead of an edge, we're talking about a very skinny line that was transmitting on an otherwise opaque background, and then by Fourier transform, which nowadays is very straightforward to execute, we get this modulation transfer function, and I show two curves in this plot.

The red curve, the upper curve, is the ideal shape. So again this tells us how the optical system for a given spatial frequency, what will be the reduction of modulation or the reduction of contrast by the optical system.

The blue curve represents the actual performance of an optical system that's being tested. If it were ideal, the blue curve would basically hug the red curve from below, but because there are aberrations, this blue curve falls off more rapidly.

So this way of testing an optical system is embodied in another ISO standard, 12233:2000, and it's essentially exactly the same technique I just described to you. It's very precisely described in the standard and can be used in any optical system certainly including microscopes.

The final point here is that the MTF again is a function of position in the field of view, and so to get a complete picture, you really should measure it at multiple occasions, not just, for example, at the center of the field of view.

Okay. One other aberration I want to mention, field curvature. More important for digital imaging than it is for viewing by eye mainly because the human eye can accommodate some degree of defocus, and just to give you an idea, and another two factors that help us out in the situation, tissue, density of finite thickness and, for example, liquid cell preparations also can be expected if some finite thickness. So the fact that the field of view of a microscope is not a perfect plane, has some departure from that, is accommodated to some degree by the type of samples that we are interested in looking at, and if we're talking about visual use of a microscope, then, of course, the properties of the human eye also help us to make the situation

with more forgiving.

This effect can be tested very readily by means of another grating object. So you can kind of tell gratings are a very common theme in testing of optics. And essentially what we do here is rely on the fact that this modulation transfer function I was telling you about a moment ago will degrade as a result of defocus and will degrade the fastest roughly of one-half of this so-called United States of frequency of the optics. So if you put in a grating with just the right frequency, you will be very sensitive to small amounts of defocus that may be present in the optical system. Again, this is an effect that's more important in the case of digital imaging where we just have a flat sensor and any defocus will cause blurring of the digital image.

So let me skip through this. Okay. One other point I wanted to make about continuity of images, in light microscopy, of course, as we move around the slide, the operator, the pathologist can adjust fine focus very quickly with the appropriate knob on the microscope.

In the digital system, there has to be some equivalent function embedded in order to accomplish the same thing, but one of the questions that I saw was going to be discussed I think in our two days here is what kinds of alignment you should look at when putting together multiple images to create one large composite, and here I just want to highlight the fact that not only should we be paying attention to alignment in the horizontal or vertical direction, but as this pair of images shows you, you also need to consider the

vertical or the axial alignment.

In this particular case, for example, the part of the image at the very top is much better aligned in the vertical directions. There are -- features. So there's a problem with the horizontal alignment, but they're both in roughly equal focus. As we move down towards the bottom of the image, where the second set of arrows is located, we find that one of the features goes progressively more out of focus than the other. So we have, as illustrated in the little inset there, the two image planes basically begin to deviate from one another. This, by the way, is a test, speaking from some experience, that you really want to automate because there will be many boundaries like this in an image, and to evaluate all of them by hand, manually, is an extremely tedious task.

Okay. So just to then start wrapping things up, what can we say about the image as presented by the microscope, light microscope to a pathologist?

So the first comparison might involve the image that's formed by the light microscope. Again, as we said earlier, it's an image that's projected to a comfortable viewing distance so that the eye is in its relaxed configuration. That might mean that the image is anywhere from 1 meter to infinity away from the viewer. It's a 3 dimensional image. If the specimen is some thickness, there is some depth to the image as well. The image will subtend on the order of let's say 45 degrees. I apologize. I got carried away

with the second significant digit there, but it's about a 45-degree full field of view. That's where we would be able to see through the eyepieces on a microscope. Of course, you have to remember that the full field of view is about 5 degrees or less. So most of that field of view is I guess being viewed with peripheral vision by a human observer.

One of the things to point out is that the sampling in a light image, and the kind of image produced by a light microscope, is going to be uniform in both directions, instead of a feature of the optics unless they're very poorly designed. The image produced by a light microscope, while it may introduce noise, if you will, due to dust or contamination or some kind of misalignment, but apart from that, it doesn't introduce noise like you might be getting, for example, from an image sensor. So an image sensor, digital sensor, will introduce its own kind of noise that is different from what you would get in a light microscope.

The light microscope doesn't degrade or diminish the dynamic range unless there's some scattering in the system, and so again that's a difference between what a light microscope can present to the viewer versus a digital camera and an image on a monitor.

And then you can certainly expect two more things. One is that the image will be very repeatable if you don't change anything on the microscope, look at the same place on the slide, the same field. Sometime apart you will see essentially the same image, and the other point to make is

that images are viewed with very low ambient background illumination. Your eyes are very close to the microscope eyepieces. The pupils are matched, and you are actually looking just into kind of a dark space at the image form through the eyepiece.

Okay. So with these two view graphs, these are really meant more for information rather than necessarily to make any particular specific point. One of the things I thought was interesting is just to kind of highlight the band of luminance levels that we can expect of an image in a microscope in terms of units of candelas per square meter, which are common luminance units. What you may expect is luminance of that image in a range of 100 to 1,000. Typical LCD panel and modern LCD panel is grated at about 300 candelas per square meter. So that kind of gives you an idea of what the comparable brightness of an image seen through a microscope versus seen on a monitor.

The pupil size in that case, the eye pupil will be about 2 to 3 millimeters. That will affect the resolving capability of the eye, what kind of contrast we perceive, and then it turns out that for looking at -- for evaluating human vision, there are several tasks that can be performed.

I'll just say a few quick words about them here. Again, the blue band in this case indicates the same range of luminance that we were just talking about, and you may hear a figure of 1 minute of arc, having to do with the acuity of human vision. That's a pretty reasonable sort of average

parameter, but there's some other levels or types of acuity that may be relevant to consider as well, again speaking of objective testing as opposed to task testing of an instrument, and that might involve things like detecting either lines or spots, very subresolution features, recognizing letters on a chart. Everyone knows, of course, the Snellen eye chart. Resolving features like you would do with a U.S. Air Force target and then finally something called Vernier acuity, which allows us to discern very slight displacements between features in an image. So each of these have their own level of angular acuity that the human eye can detect, and they also depend on the contrast of the features that we might be viewing. So that's sort of a way to evaluate microscopes or digital systems taking into account the performance of the human eye and the human visual system.

Okay. The final view graph I want to show you has to do with kind of the situation having to do with standards in light microscopy, and the news I'm afraid is bad. There really aren't very many standards especially for light microscopes and what we might expect of their performance. There are a few things scattered around, however, that might be relevant, and I'll just mention some of them here.

The first one is this ISO standard I mentioned a short while ago, the slanted edge testing of the performance of a microscope. There is another standard related to still-picture cameras that has to do with evaluating how a camera converts variations in image luminance to variations

in digital signal, how accurate that is, how linear that may be. There is a standard that has to do with evaluating scanners that digitize pieces of photographic film, and that defines a number of parameters related to resolution and other factors involved in these kinds of instruments. Again, in some instance similar because we're talking about an image on film versus basically a tissue section on a glass slide.

And then the last two standards I mention here actually are directly related to microscopes. They're fairly recent, and they define what ratings on a microscope objective related to either flatness of field or absence of field curvature put differently, or correction of chromatic aberrations, what those would translate into in terms of actual physical performance.

So there are a few standards, but it's at the moment, light microscopes have been around for 400 years, probably much longer than anyone really thought about developing standards, and I think we're in some sense are seeing a consequence of it here.

That's all I have. Thank you.

DR. ADCOCK: Thank you. I'd like to thank Dr. Descour for his presentation.

Does anyone on the Panel have questions for Dr. Descour at this time?

DR. HEWITT: Yes, ma'am.

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: Dr. Descour, I have a couple of quick questions. You talk about illumination and Kohler illumination with the use of a filament light source, but you don't mention anything about the use of either a light pipe type of light source or rather diffused but intense source or an LED light source. Could you make some comments about how that illumination alters the mechanics of image capture and/or contrast? And then just as a secondary, second small question, what are the effects of the introduction of ancillary optics, primarily doublers or teleconverter-like objects within the light path with regards to again contrast and resolution and depth of field?

DR. DESCOUR: Okay. Thank you. Well, as you point out, there are several options for illuminating the specimen in a microscope. LEDs are certainly a very up and coming kind of an option. I guess I'll answer the question right now in the following way, that for image quality to be maximized in the light microscope, at the specimen, we are looking for two key properties of the illumination. The first one is that the illumination be uniform across the field of view, and by uniform I would mean just the light level across the specimen seen by any point in the specimen is the same. And the other key parameter is that the numerical aperture of the illumination also not vary across the field of view.

As we mentioned, closing or opening the aperture diaphragm can affect the appearance of the image contrast resolution tradeoff, and so to get a consistent image, the numerical aperture has to be the same across the

entire field of view.

How that relates to your question is as follows, that if one uses a diffuser, for example, that is very likely to satisfy both of those requirements at the expense of efficiency. So, you know, from an engineering standpoint, a brighter light source may need to be employed, but you will still be able to get a uniform illumination level. The numerical aperture is likely to be controlled by the properties of the diffuser and may be more difficult to adjust than you would be able to on a regular light microscope where you have an aperture that accomplishes that. It's likely to be controlled by the properties of the diffuser and may be more difficult to adjust than you would be able to on a regular light microscope where you have an aperture that accomplishes that.

In the case of some kind of a light pipe, that's also a good way to scramble or uniformize the appearance of a light source, and it may allow us to replace some of the optics in the illumination system but again the key issue there would be to make sure that it's not only that the illumination is uniform spatially, but that it also has the same angular characteristics or numerical aperture across the whole field of view. So that would be the answer to the first part of the question.

Now, as far as adding more optics to the optical system, the most common example I'm familiar with is putting in the lens that will essentially double the magnification of the optics, and I think here probably

the most important issue would be that, and I didn't spend much, in fact, I didn't talk about it at all in this talk, and that's the idea of empty magnification, I think it would be relevant here, that the microscope objective, as you may recall, has the capability to allow us to resolve certain details, and that's limited by its numerical aperture and some wavelength in the visible part of the spectrum. If an image is magnified beyond what the capability of the microscope is inherently, all we're going to get will be a larger image but one that doesn't show more detail.

Now, that may be more comfortable for viewing by eye. You're not asking the observer to look at the image, sort of a limit of resolution, but it may not introduce any more information necessarily than is ultimately limited by the capabilities of the microscope.

So the depth of field, the resolution, those would really be governed by the microscope objective. The addition of some doubling optics or additional magnifying optics may play a useful role if the numerical aperture of the microscope is high enough or may simply enlarge the image without showing more detail.

DR. ADCOCK: Dr. Birdsong.

DR. BIRDSONG: One of your bullet points a few slides ago, you mentioned that the optics don't reduce dynamic range, and I guess maybe this is a little outside the scope of your talk, but with regard to histologic preparations, do you have any data on what the dynamic range is, you know,

in other words, what would a digital system need to accommodate to have the full dynamic range of a histologic preparation?

DR. DESCOUR: I think that's a great question. I hope that some of the other speakers may address it as well, you know, it certainly has to do with how the tissue is stained. So it kind of takes --

DR. BIRDSONG: Right, H&E stain, yeah.

DR. DESCOUR: You know, most of the systems, the additional systems that are most common will have, most commonly will do imaging in three colors, and then it's going to be 256 different gray levels per color. There is likely to be more variation in a specimen, but again that's just going to depend on what kind of staining was done. So that's a little harder to answer just because it now involves not just the instrument itself, but what it is that you're looking at.

DR. ADCOCK: Dr. Kulesza.

DR. KULESZA: You talked about the slanted edge testing. My understanding is that there is no Z to that. There is no Z component in that testing, and then you transitioned to the transfer function for defocus, right, but in the specimen you have a thickness that's already there whereas looking at defocus, it's still departing from an infinitely thin specimen. So are there tests, perhaps somewhat analogous to that defocus test, that you can think of that would look actually at a thick specimen say of a 5 micron standard and look through that specimen and tell me how it's focusing in that Z of 5

microns say, arbitrarily assigned as 5 microns.

DR. DESCOUR: Well, the answer is most tests unfortunately are biased towards an infinitely thin object. So it could be a metal coated mass with some patterns on it like we were talking about. I think tomorrow you will hear about some test targets that do have some 3-dimensional structure to them, biological targets on a glass slide, that could be used to perhaps even objectively evaluate how well a system does with an object of infinite thickness, but all the tests I'm familiar with for testing an optical system in terms of its capability to resolve and contrast and so on basically make the idealized assumptions that we live in a very 2-dimensional world. So that's perhaps an opportunity to develop some specialized targets that would be relevant to this field that would capture the third dimension that you referred to.

DR. ADCOCK: Dr. Zhou.

DR. ZHOU: So I tried to tie your theory with the statistical theory. So I was looking at -- so you have actually introduced several graphs and some formulas on how to turn the imaging to numbers so you can draw the graph. If I understand it correctly, those numbers actually derived under certain model assumptions like perimetry models. For example, on your page you show the two curves, two distinct curves merging into the same curve.

So what type of perimetry model in the optical evaluation used, and then will those assumptions make any sense or will it depend on which

particular applications, on whether the assumptions are reasonable because that has the implication down the road when we try to evaluate our device here. Are you aware of it?

DR. DESCOUR: Well, I think the most important assumption that is behind the resolution testing methods I described to you is, and these ideas like the Rayleigh resolution criterion, for example, is the absence of noise in the images. In the presence of noise, you know, the Rayleigh criterion, depending, for example, on the light level, and this may take us more into fluorescence microscopy than bright field microscopy, you may find that there simply aren't enough photons being detected by a sensor in order to confidently say that we see a dip between the two peaks which would then signal the presence of two point sources close to one another versus just one larger source.

I think, you know, to a large extent, to the best of my knowledge, when testing microscope performance or optical system performance, the issue of noise and how they might affect these idealized assumptions tends not to come up that often because the images are taken at illumination levels where noise is just not that significant. So, you know, we can talk about the Rayleigh criterion, if signal-to-noise ratio is high enough that the presence of noise in the image state, a Poisson or some other noise generated by the sensor itself, is not a big issue.

DR. ZHOU: Do you have to assume a Gaussian process or

Gaussian distribution for the curve, the two curves you show us?

DR. DESCOUR: No, those -- so the diffraction patterns that we were talking about that are sort of behind the resolution criteria, those just have to do with the way that light propagates, the waves propagate when they're focused and how tightly you can produce a focal spot. So in that sense, that's a deterministic process the way we typically treat it.

DR. ADCOCK: Dr. Mello-Thoms.

DR. MELLO-THOMS: You mentioned in one of your slides that the luminance range for the light microscope goes between 10 to the 2, to 10 to the 3 candelas parameter, and the LCDs is about 330. That would put it around the -- saturation zone on your graph. In radiology, we've observed significant changes from moving radiologists to look at film to look at the digital images. Are you aware of any related studies that have been done in pathology?

DR. DESCOUR: No. I think there's -- the observation of how pathologists interacts with an imaging seen through a light microscope, to the best of my knowledge, I can think of only one study that was done by Elizabeth Krupinski at the University of Arizona, and she just reported to me a few days ago that there was a paper given by somebody from UPMC at a recent conference, but I think, you know, that perhaps it's kind of a difficult task to monitor how pathologists might look at images and what light level they might typically operate, I think that's kind of an undiscovered country,

and it hasn't been studied very extensively and quantitatively thus far to the best of my knowledge.

DR. ADCOCK: We'll try to take two more questions. Dr. Gilbertson and then Dr. Hewitt.

DR. GILBERTSON: Clearly what you've noticed, as the user, you can manipulate the image directly, you know, in addition to changing the quality directly. The effect of this, I mean in terms of how in a normal use of a microscope, these changes, are they greater than or less than the changes you would normally see from, for example, change in histology which is, you know, another part of the microscope system obviously. Is it a bigger problem or a worse problem if you can't control that?

DR. DESCOUR: I guess I would say that's a great question. I think, you know, the microscope is set up correctly, I would guess, and it is just that, a somewhat educated guess, that the variations of specimen and staining quality might overcome, you know, whatever range of adjustments that a pathologist might make to a microscope. But it is only that, and I'm not aware of any formal studies that would look at, you know, what would be the effect of the variations of the specimen staining versus adjusting the microscope to variations on the aperture diaphragm, for example, and how they would affect the ultimate outcome, which would be the diagnostic performance by the pathologist.

DR. GILBERTSON: But one particular thing is the optical -- the

depth of field. Clearly it's -- given a specimen of let's say 6 microns and an objective lens, you're not even sure if it's a, you know, 20X lens or 40X lens, how many depth -- how many sections will you cut through a 6 micron section at 40X, a typical 40X lens?

DR. DESCOUR: Well, here, if you'll permit me to get on my soapbox, to answer that question, the parameter that needs to be kept firmly in mind is the numerical aperture of the microscope objective, which tends to be correlated with the magnification. In that case, I would estimate probably three to four images of the outside would be needed in the vertical direction in order to give you a view of sort of equal increments through a section of that thickness. As the numerical aperture increases, however, say to 40X, .95 numerical aperture, something like that, that number will grow because the depth of field will get skinnier and skinnier to a submicron.

DR. GILBERTSON: Great talk. I have many questions, but I'll stop. Thank you.

DR. HEWITT: Okay. Two more questions unfortunately. Back to the refraction gradients. Those gradients don't usually have a cover slip on them. Is the focal point, when you're creating an image in the presence of a cover slip on the cover slip, in the middle of the tissue or on the surface on which the tissue sits? That's the first question.

The second question may be harder or easier than that. With issues of noise and resolution, what about the effects of wavelength both to

the light input into the specimen as well as the wavelengths that can be detected by the detector, some detectors being able to reach outside the visible range of light that the human eye can perceive? Will that introduce an alteration in the calculated resolution of the system?

DR. DESCOUR: I'm going to have to extrapolate on the second part of your question. I think you could expect that. A silicon detector, for example, that's properly filtered could see farther into the near infrared than the human eye can, and as we talked in terms of this minimum spot size to which light can be focused, that will grow as the wavelength increases. So unless filters are present in the system, you could, I think, expect some reduction in image quality as seen through a silicon sensor versus a human eye. That's a problem that can be fixed relatively readily by introducing filters into the system that mimic the spectral response to the human eye, and I'm doing my best to forget the first part of your question here.

DR. HEWITT: Cover slips and focal planes with relationship to the gratings that typically lack a cover slip. They are usually just a -- on glass or mirror surfaces.

DR. DESCOUR: Oh, yes, that's an excellent point. That's right. I didn't really touch on that, but that's, yes, that's a very important issue. So when evaluating the performance of microscope optics, they were designed with a coverslip in mind. You know, I told you that one could prepare certain kinds of targets that would be used to kind of test the performance of the

optics. Those would have to be cover slips. I think that's what you're asking. You would have to recreate the conditions under which the microscope objective is meant to perform, and so if it's designed to have a 170 micron coverslip, that would have to be on top of the resolution target.

DR. HEWITT: That does answer, but I guess I'm getting actually to a different question, and maybe it wasn't clear, and that is, let's say you have a lumpy, bumpy specimen. And, of course, it has cover slip media that's covering it to make it flat and that glass is applied on top. Where's the focal plane when I make an image? Is it actually the bottom of the cover slip, or is it somewhere in the tissue, or is it the glass on which the tissue sits?

DR. DESCOUR: I would expect that it would be somewhere in the tissue. The microscope objective is typically designed assuming that the specimen is right up against the bottom of the cover slip, but I think in practice you can expect that that will not be the scenario under all circumstances. So the object plane would actually be recessed a little bit into the mounting medium below the top of the cover slip. The index of refraction, I think, of the cover slip, the mounting medium can be expected to be very close, so that shouldn't introduce much aberration, but this would be somewhat of an off-design use of the microscope if you will.

DR. ADCOCK: Thank you, Dr. Descour.

DR. DESCOUR: Thank you.

DR. ADCOCK: I hope that we may have some time later in the

day to ask questions of our guest speakers again.

At this time, I'd like to have our next guest speaker, Dr. Ulysses Balis, an Associate Professor of Pathology and Director of the Division of Pathology Informatics at the University of Michigan Health Systems, come to the podium.

Dr. Balis will be speaking on the features essential for interpretation of histology and cytology glass slides.

DR. BALIS: I'd like to thank the FDA for the opportunity to comment this morning on this topic. I'm aware I'm starting a little bit late and should be able to catch us up in time for questions.

So my presentation, I think you'll find, complements the prior one in that it takes the information which has already been presented from a technical perspective of optics, and specifically optics applied to microscopy, and then looks at that from the specific prism of identifying the features that are essential for diagnosis.

So it was stated earlier this morning that the intent of this Panel is to identify global attributes or qualities of whole slide imaging systems and not specifically organ specific or diagnostic specific attributes, and therefore in looking at this title, I think it's important to provide just some qualification that when we speak of features essential for diagnosis, the features are really optical system property features, not diagnostic features, and thus my presentation really focuses on the appropriate or possible appropriate use or

selection of optical properties in digital system properties in concert with optics to allow for the broadest and most effective representation images that have the highest diagnostic accuracy and sensitivity. So I will not be speaking specifically of diagnostic features of diagnostic entities, but obviously some histology and cytology examples do come into play here.

So as an outline, I think it's important to provide a very brief historical context of optical microscopy, and fortunately I don't have to go through this in too great of detail with Dr. Descour's very helpful overview already.

I would like to point out some specific limitations in optical systems that tie in to the selection of features and then continue with an overview of those histologic feature classes that are essential for diagnosis, again not in an organ or a specific diagnosis category orientation, but from the perspective of the interaction of light and the H&E stain and the images that are rendered on the digital imager.

Then that's a good segue for going into the whole slide imaging platform itself and this interesting new paradigm presented to us, which is this optical, digital hybrid and what that constitutes in terms of opportunities and risks for accurate and consistent use of this platform for diagnosis, and that ties in directly to the limitations in the current digital representation model we have, and then the potential real world consequences of those limitations.

And then finally some simple use cases which I think highlight the current state of the art limitations of how whole slide imaging can in some instances subsample or provide images that fall short of the current gold standard, which is optical microscopy, and then finally some closing remarks.

So, in effect, microscopy as defined by the use of optics has several hundred years of compound microscope experience, in fact, it is nearly 400 years old. It's a well-established technology and has been stable now for 100 years from the perspective of the use of compound optics. Dr. Descour alluded to this but, of course, with the need for correction of chromatic aberrations and color issues within optics, the observation and studies of Abbe in the 19th century, the advent of so-called crown and flint glasses which have different refractive indices allowing for concurrent compensation of refractive artifacts leading to significant improvement in color quality, and this is a feature of modern microscopes that we've enjoyed for practically 100 years now.

More recently, microscopy has undergone a number of incremental innovations in the perspective of better quality coatings, better numerical modeling of the objective elements to increase both contrast and total light gathering capability and field of view, again multiple parameters that can be optimized concurrently, leading to the best possible viewing experience. And it's important to note that these are psychometric qualities that are in some instances difficult to provide quantitative measures but, as

pathologists will look at various microscopes and images, usually can come to agreement that one image is superior to another.

And also historically, microscopes have been exempted from the rigorous certification that we currently observe for other radiology type and other medical imaging modalities which have very specific metrics of performance, and also the appropriate functioning of microscope equipment is, I would say, based on what we could call the guild model, where the artisan is expected to know their tools and know when their tools are functioning properly and know when they're in need of repair.

So in characterizing optical properties of the microscope, there are issues which have already been touched upon: magnification, diffraction-limited capability, that is to say having the optics or objectives in place to capture all the spatial information that is theoretically possible for a light gathering system, and in terms of diffraction-limited capability, there is often mention of the ability to represent the image at the point spread function or limit, again consistent with what's already been presented to you from Dr. Descour. So I won't have to elaborate on that.

And then three very important areas, contrast or contrast and brightness, tie in then to field of view, numerical aperture, and color correction, and then finally the spatial aberrations and the lighting and uniformity issues. And of these, as you'll see, the whole slide imaging imparts an interesting opportunity at the same time as imparting a challenge, which is

with the ability to capture the entire digital dataset of the whole slide. You now have the ability to create virtual images which are subsample. So, in essence, a perfect low magnification image, which is far better than could be obtained by optics, by the process of taking the high resolution or high spatially sampled image and numerically subsampling it, creating a virtual magnification image, which theoretically has perfect optical properties, and then you're limited in quality to the characteristics of the display system. So at that point you have transcended the characteristics of an optical system, and now your performance is predicated outside optics, which is in the realm of digital display, and this is really where the whole slide imaging has challenges, and now you've imparted an additional section to the overall pipeline of processing leading to the ultimate rendered image on a display.

If we look at contemporary microscopy in surgical pathology settings, we really have no official regulatory oversight of the minimum contrast in resolving capabilities of the microscope equipment itself, and in discussions I've had with pathologists over the past month leading up to this Panel discussion, I was able to confirm to myself that, in many community-based settings, microscopes are being used which are probably are well past their retirement age, that is to say, 40 or 50 years old, and in some cases with only 2 or 3 objectives as opposed to the normal 4. So there's no guarantee at all that in every setting, the same quality of microscope will be used for similar types of diagnostic challenges. So with this as a baseline, I would just

offer as a high level concern, that not only is there an opportunity to provide rigor in the certification of performance of whole slide imaging, but similarly in microscopy itself. I think this is something which has gone too long perhaps in the guild fashion.

We must also consider the suitability of the modality for consistent diagnosis and performance with various classes of subject matter. Again, we police ourselves as a specialty reasonably well. For example, most hematopathology settings make use of the -- objective, but I don't know for a fact that that is a universal reality.

We have at present no agreed-upon minimum set of functional capabilities for the ergonomic interface itself, knowing, of course, that a poorly designed interface can create a risk factor, and if you look at a metric, studies about 10 years ago from the Anesthesia Safety Foundation, they identified a checklist of questions that an anesthesiologist could ask themselves before participating in an operative procedure, that could, in fact, minimize operative risks by over 300 percent, and it turns out that this checklist involves better designed equipment that allows rapid identification of the correct functioning of that equipment. And similarly, microscopes at this point have no standard design guidelines. There are similar features, which are not by not any stretch constrained to a universal standard of performance.

And finally, which is perhaps the greatest concern, there is no

initial or periodic competency testing of the user's ability to actually calibrate the microscope to its theoretical resolution limit, so-called Kohler illumination. And again, an informal questioning of a number of colleagues who use microscopes confirmed at least for me that the ability to carry out Kohler illumination is not universal. And so at baseline, not only is there concern about the use of a microscopy platform as a consistent technology, but there is also concern about the ability to use that platform to its maximum capability.

So I can go over the inherent limitations in optical systems quickly as these have been covered. Very briefly, I would want to point out that in the relationship of diffraction limitations of light and the interaction of the subject matter, which was the slide, an interesting question was already raised about the magnitude, order of magnitude of the -- caused by the tissue and the variability of staining upon that tissue and the variability of the optical systems themselves, the quality of the objective. And in point of fact, these are different types of artifacts or potential artifacts. For example, from the quality of the section in the staining is going to affect the dynamic range and the overall offset. So these are basically indices of light level, the number of photons that can be gathered versus the optical properties of an objective are going to lead to its varying point spread function or the physical resolution.

So, resolution and dynamic range are both important in

ultimately rendering an image which is close to the actual reality of the information on the slide. So neither of them can be discounted. I think they need to be measured independently towards the goal of characterizing a system.

And then finally in terms of overall calibration of a microscope, again as was touched upon, there's a need for uniformity of the light source so that features within the light source itself, such as filament do not modulate with the structural components of the histologic section to create a compound image which would potentially be more difficult to interpret.

And then color temperature can affect the consistency of practice since, in fact, we are talking about color subject matter, and if the color temperature is not uniform, that could potentially bias the interpretation of the histology.

And finally, field of view is an additional metric which is a tremendous opportunity of whole slide imaging which was not previously easily addressed in the setting of optics, which had constrained pupil eye spots and light gathering angles as defined by their intrinsic numerical apertures. Numerical aperture, in essence, as a concept goes away with whole slide imaging, and this an important point that I'll touch upon in a moment, in the advent of being able to carry out post-processing of viewing operations on varying degrees of larger and larger screens. So this is actually one of the strong points and very exciting aspects of whole slide imaging, and

that field of view now is no longer limited to the constraints of the physics of generating optics that are suitably large in their light angle collection.

And finally I'll just mention that ergonomics design is important. In looking at the essential histologic features, we can divide this down into a number of hypotheses. Length scale is important. Pattern recognition is important, and this ties into the practice and the manner in which pathologists use histologic imagery to render a diagnosis, and that's going to be applied for histology and also for cytology.

So firstly, length scale issues, pathology diagnosis by microscopy is an interesting art form, if you would, or execution of steps which typically falls along the multiple length scale hypothesis. Initial assessment is typically carried out at low magnification, low power objectives, to obtain architectural pattern recognition, architectural features. Then use of successfully higher magnifications is employed to confirm features that architecturally might be associated with specific diagnoses that can then have cytologic features confirmed at the higher magnifications. And so the typical progression of rendering a diagnosis starts from the lowest power and progresses to successive higher powers until the highest magnification is used.

Which highest magnification is used is in many ways predicated on the subject matter, with the far majority of surgical pathology cases being comfortable within the 40X objective or 400X total system magnification but

hematopathology requiring oil immersion or equivalent of 1,000X total system magnification.

Recognizing that pattern recognition is really a low power histologic architecture step followed by successively higher resolutions, there is the need to recognize that different objectives are optimized for different particular use cases. So the low power objective is typically going to have a relatively low numerical aperture. Therefore, it has a high intrinsic depth of field, and the higher magnification will have higher magnification capabilities, more resolution, and intrinsically then because of the optics of capture, a lower depth of field, and therefore at high magnification, one is more susceptible or the applicable system is more susceptible to perturbations in the selection of the focal plane, and this is extremely important when looking at the transposition of the capture model to whole slide imaging because typically whole slide imaging will generate the sum total subject matter at a 20X or 40X objective, which is using a higher numerical aperture device, which means therefore that the entire surface area is much more sensitive to perturbations from deviation from the optimal focal plane, whereas if one were to capture at the low magnification setting which has an inherently low numerical aperture and high depth of field, there would not be this concern.

So the caveat of being able to generate a wide-field, numerically subsampled image, which is a low power image from the whole slide image, while you have the high field of view or very broad field of view,

you may be more susceptible to artifacts of focal plane, and this, of course, plays on the reality that the tissue section in histology is not universally flat, and in specific cases of cytologic feature classification at high magnification and low magnification, this is compounded by the reality that various structures have inherent 3-dimensional structure, and when sampled with high numerical aperture, you are only looking at a very narrow focal plane of that thicker section, and that is a limitation of data acquisition.

So, in summary, microscopy-based diagnosis is a stepwise iteration of discovery through low magnification to progressively higher magnification.

So in histologic structures, we have larger features which can be identified at low magnification, that span multiple cells or many groups of cells, and there's high variability in architecture for many diagnostic entities. And so this is the step at which pathologists are forming their high level opinion of a diagnosis, based not on cytologic features but again on architectural features which in many instances are not as susceptible to having the presence of extremely high resolution. Again from a Fourier transform frequency theory perspective, the large majority of information which could be considered of diagnostic quality here is low frequency information. So given that, perhaps the fact that information is scanned at high numerical aperture with the potential out of focus capture is not so important if, in fact, the diagnostic elements present are low frequency

elements.

As we proceed to progressively higher magnification, and so that's cell pattern diagnosis at 20 and 40X or perhaps also 100X objective, we are now looking at glandular structures, we're looking at cytologic detail 3 dimensional cytologic detail. There's significant need for depth of field, specifically in cytopathology where structures can have real depth that can span as much as 15 microns, therefore requiring, if one were to exhaustively sample all those 15 microns at a high numerical aperture objective, as many as 40 sections. We're clearly not sampling typically in whole slide imaging 40 sections for the entire whole slide image. There are technologies that will locally sample at higher spatial periodicity but not the entire surface area, at least not yet.

So if we look at the combined reality of whole slide imaging, and the appended need for post-processing, which is in the digital dataflow realm, we have a combined set of artifacts which is new and potentially introduces new risks for data degradation or diagnostic interpretation degradation.

We have, of course, the light source. We have the section itself, which is transilluminated, and then we have the capture of the microscopic image. And this is the first step which is new. It's typically some type of silicon-based sensor, either for TDI or linear or grid based stitching, and then a number of shuttle steps where that digital information has to be

transported through various stages in the rendering system, both storage and then re-rendering. And then finally, after file storage is the re-rendering of the image on a display system.

So if you look at the total system quality of this new construct which has a front end, which is conventional optics, followed by digital processing, there are more steps in which error can be introduced.

If we look at the limitations in digital representation of optical imagery, there are still opportunities for degradation of image capture at the light end. And, of course, this ties in mostly to resolution itself, and then contrast and brightness.

So for example, with the rendering of Kohler illumination, there's actually an additional setting besides the substation condenser, there's the field diaphragm, and that limits the amount of light which is illuminating the overall culminated pathway through the entire system.

The reason this is important is not as much for resolution but for stray light. So with excess light going into a system, that has a net effect of reducing contrast. So optical systems, because of their inherent need for alignment, when combined with digital capture, become exquisitely sensitive to stray light for the simple reason that typical digital capture systems, vis-à-vis CCD chips for example, have a finite dynamic range by virtue of their electrical design. If you exceed the light levels, you end up saturating the system, and stray light has the effect of effectively diminishing your dynamic

range. So the current technologies which are based on typically three channels of 8 bits per channel fall short of the actual inherent dynamic range present in the light column, which can be as much as 35 decibels beyond that, or in terms of numerics, we're sampling at 8 bits. There's 0 to 255 shades of gray, when we should be sampling in at least 0 to 4,096 or 0 to 8,192; in other words, the need for 12, 13, or even 14 bit sampling. So this is an opportunity to address that limitation by increasing dynamic range, and that would be dynamic range for each color channel.

An additional problem with the fact that we've appended now to this data model, a digital set of rendering tools or digital technology, we have now transcended this artisan model where the pathologist is not really trained in effective understanding or effective utilization of digital dataflow models or understanding the physics of a CCD device, and therefore when artifacts are introduced into a system, it's not clear, at least to me, that the pathologist or the interpreter will be qualified or able to identify at what stage an artifact potentially was introduced or what artifact, in fact, has been introduced, and this, I think, is an area for concern.

And then two conclusions or possible conclusions, it follows that adequate dynamic range and spatial resolution are essential attributes for enabling digital images to be diagnostically equivalent to optical microscopy.

And it also follows that the same absolute metrics of whole

slide imaging performance would be helpful in compensating for the inherent lack of an artisan model. So if you look, for example, at engineering requirements in ultra high reliability models like aviation, there is the notion that every part which is a mission critical part of a factor of safety. And so in recognizing that there are transitions in image quality from what would be considered equivalent or superb image quality which is diagnostic to inadequate image quality which could lead to an incorrect diagnosis, I think we can all agree conceptually that there is some transition point at which an image is rendered unusable. But that image point is potentially difficult to discover.

So the notion of a factor of safety allows for the operation of a combined hybrid optical digital system in a parameter space which hopefully lets one run independent of running into risky or dangerous parameter spaces.

So if we start tying this all together on the effect of the real world consequences of the fact that we have an additional set of digital steps now on optical microscopy and how this affects interpretative accuracy, there's a length scale issue that's always at play. In other words, you want to have the best possible image quality at all magnifications. And again, at low magnification, the numerical aperture could generate sufficient field of view but inadequate spatial resolution if it's too low. Fortunately, we don't have this problem in whole slide imaging because typically the numerical aperture

is from a high numerical aperture, higher magnification image. This is actually a great opportunity, as I mentioned earlier.

So the problem, of course, with this model is that typically with high numerical aperture systems, the field of view is smaller because it's traditionally looking through a microscope. Herein lies the opportunity and challenge at the same time. It's not so much a question in the optical domain of the numerical aperture. That is a given that it's going to be high. It is what is the field of view rendered to the pathologist's retina, and this really has two direct variables.

What is the size of the rendering engine, the visual panel that the pathologist is using? Is it a whole wall? Is it a tiny little monitor? And how far back physically is the pathologist from that view screen, which again affects the actual rendered resolution on the retina?

So part of the standard potentially involves not just the selection of the rendering display technology, but also its physical size and the recommended distance that the pathologist should be from that viewing plane, in order to get the actually calculable rendered resolution on the retina. So I would just offer that these are parameters that should be factored into the characterization of system performance in addition to just the digitizer, the CCD element itself.

And if we look at real world use cases where whole slide imaging modality may not be equivalent to optical systems, and why this is so,

in hematopathology, we have resolution-based concerns in the setting of relatively uncommon oil-based or oil immersion whole slide scanners. So without those being in play, you're looking typically at a 40X maximum acquisition, which may be inadequate.

And in pigmented lesions, where the dynamic range of the actual intrinsic pigment within H&E sections is beyond the hematoxylin and eosin, exceeds the 255 shades of gray, and thus the pathologist or interpreter is left looking at what is effectively a black splotch on their screen which is diagnostically not helpful. And again, going back to the notion of the factor of safety, now that image has entered into this potentially dangerous transition zone between diagnostically acceptable and diagnostically unacceptable, and then it becomes testing of a skill set of whether or not the interpreter, or pathologist, cytopathologist, or histopathologist, has the appropriate clinical human skills to know whether to defer to light microscopy. And so this creates a challenge. Is the technology acceptable, the whole slide imaging technology acceptable in current form for every possible use case, and at present, with 255 shades of gray for each color channel, I think that question needs to be very carefully explored so that a cogent set of use cases perhaps can be excluded from routine diagnosis until both the resolution and dynamic range are improved.

And in closing remarks, and actually there are a number of slides, I think that I loaded it wrong, this is not the right presentation, but I

can show the images afterwards. There's time for questions.

Given that optical microscopy works remarkably well as a diagnostic tool despite its relative lack of standardization which I think is good news, there is inherent robustness in the human visual diagnostic interpretive process, and we would expect this robustness to carry through in the continued use of whole slide imaging, that is to say for interpreters to recognize when an image is falling short of being diagnostic.

However, I think there are still opportunities to provide formalized direction to where we would expect that functional degradation is more likely to take place, so interpreters can either steer clear of entire subject matter classes, or if compelled to use them for medically required situations, emergent situations, to approach those diagnostic settings with extreme caution. For example, one possible use case could be the necessity of carrying out diagnosis of a pigmented lesion via a remote review of a whole slide image for medical necessity because a pathologist is not present in a remote site and which might be one possible use case.

And finally as whole slide imaging continues to make strides in resolution, depth of field, dynamic range, which I think that with each successive offering of a new generation, that is an opportunity for re-characterization of that system in the conventional models which have worked very well for us in the past, which are the receiver operator characteristic or ROC curve to optimize sensitivity and specificity and

maximize the chance of carrying out an accurate diagnosis without a false positive rate. And I think the field at large can then benefit from the use of this technology in incrementally more use cases.

And then finally, the field of pathology at large, being ignorant of performance characteristics of the digital aspects of these optical systems, needs to benefit from some level of external adjudication on the baseline application of the digital technology as appended to the optics so that minimum standards of resolution and resolution matching are appropriate for those optical systems. In point of fact, the chance for undersampling of information within an improperly designed or selected CCD, which renders equivalently a blurry image which is the first recognized problem, is equally problematic as oversampling where you generate empty magnification unless there needs to be significant care in the selection of the digital components that make up the total whole slide imaging system.

And at present, given that this is really conferred totally now to the aegis of engineering teams of each group that manufactures or designs their systems, I think that having metrics of performance that such engineering groups can look towards for direction would be very helpful. Right now there is no universal standard of minimum engineering performance of such appended digitization systems as added to the optics.

And finally, adoption of whole slide imaging technology should be validated on an organ system diagnostic entity basis for those areas where

we initially identify that there is limitation in the current dynamic resolution to be able to carry out such identification.

And this ties into an interesting opportunity and concern, the best way that we've done that so far has been kappa statistics or variability, and I think that more work needs to be done in this respect just for light microscopy itself so that we have the true gold standard for the interviewer variability of a diagnosis on a microscopic or optical platform before we then make the really extrapolatory statistics of what is the equivalency of whole slide imaging to microscopy. I think we need to understand microscopy better first.

And those are my comments. Thank you.

DR. ADCOCK: I'd like to thank Dr. Balis for his presentation.

Do any of the Panel members have questions at this time? Dr. Davey.

DR. DAVEY: Thank you for a great presentation. I think we all recognize some of the limitations with like cytopathology specimens for screening for rare events, but it occurs to me that with histology, there are going to be problems especially with large pieces of tissue with rare events, and like especially if you're doing looking at a large number of slides from a specimen for a staging procedure, for example, looking for clusters of tumor cells in a lymph node, margins, that kind of thing. So I just wondered if you have any commentary, and also, I mean I'm a little concerned about the field

of view being so, you know, large at low to mid magnifications that it'll be easy to miss rare events, where if you saw those, then you would go to the higher power and look at them and probably make a diagnosis, but that's one of the biggest concerns I have. So I just wondered if you know of any information on that.

DR. BALIS: Well, as a preface, I agree with both of those concerns. So with the first concern, which I think can be addressed of the fact that histology does not always occur in planar sections. There are dog ears. There are areas that are out of focus and, in fact, for example, let's use a prostate biopsy. Prostate biopsy is a metric of search for a rare event. Theoretically one glandular lesion could have the diagnostically important area, and if it's out of the plane of focus, one could miss that.

In that setting, if we look at the current use case, which is our gold standard of how we mitigate that, the reality is as the interpreter is moving the slide constantly, one hand is firmly on the focus wheel constantly making adjustments. So the fact that something is out of focus actually falls away from the equation. There's essentially an intrinsic human driven auto focus so that even localized areas that are temporarily out of focus are brought into focus so they can be reviewed.

There are multiple solutions to this problem from a digital capture perspective, some of which have been implemented, some of which could be implemented, that involve capturing more than one plane for a

whole slide image, so that statistically one of the planes captured could be flipped and then some type of user ergonomic interface allowing for selection of the most optimal plane, thus essentially recreating the experience of a pathologist reviewing a slide, because again if that's our gold standard of performance which works, then having the equivalency in terms of that depth of field information should mitigate that rare event which might be on a different focal plane.

Towards your other concern, which I also share, of having too much field of view, that leads to rare event detection. Again too much of anything is potentially a bad thing. Again, it's a matter of constraining it or optimizing it. So it's the appropriate field of view because again we lose resolution as our peripheral vision is extended. So I'm not saying it has to be a whole wall, but that's possible, of course, but that may not be the optimal practice modality. I think there's opportunity to explore what is the optimal rendering of size and what is the optimal distance for which the interpreter should be seated or standing away from that image. And that again can be perspectively tested and then optimized, and then select technologies and recommended rendering engines can be, with the specifications, can be based on those optimal findings.

DR. ADCOCK: Dr. Zhou.

DR. ZHOU: So I have two questions. I wonder based on your experience which reader variability is larger, current technology light

microscope or the whole slide imaging. So that's the first question. And actually the related question is which actual technology has potential to reduce reader variability? So the second question is what are the gold standards you use to compare the whole body, I mean whole slide imaging based on the pathology point of view?

DR. BALIS: Okay. So for your first question, in terms of looking at performance characteristics, I think that the studies to date have been somewhat equivocal in being able to demonstrate that whole slide imaging was appreciably lower in diagnostic equivalence to use of a microscope. One of the best ways of looking at where caveats for diagnostic detection is in, for example, studies with frozen section diagnosis. University of Pittsburgh has carried out a number of studies in looking at which classes of cases were deferred. So I think that gives us some significant insight to which cases where the information content of the whole slide image fell short, requiring final review of the glass slide to render a complete diagnosis. That's one class of evidence, and the other is then looking at the historical body of evidence from such controlled perspective studies, where there was discordance, between the initial reading from the whole slide image and the subsequent reading, when differing, on glass, what types of diagnoses.

And so I think there are models by which we can identify those differences and, in fact, those types of perspective studies already have shown that there are differences, but they're quite subtle. So the diagnostic

accuracy of whole slide imaging for general use case in histology, actually this was the challenge, falls within the noise floor, if you would, of the re-diagnosis by light microscopy itself for most diagnoses.

So a similar question asked is what is the quality of rendering of pathology diagnoses by pathologists using a microscope over again? It's not 100 percent, and recognizing that, you have really the convolution of two statistical error rates which are in the same order of magnitude. So that becomes a statistical challenge which I don't know how to answer but I hope you do.

Towards your second question --

DR. ZHOU: Actually, one additional one, the reader variabilities, which one is bigger? Is the whole slide imaging or the light microscope?

DR. BALIS: I don't know that that's been definitively answered. I certainly don't know the answer to that. My suspicion, and again this is just an extrapolation on my part, is

that the interviewer variability exceeds the noise floor imparted by the conversion to whole slide imaging. I would be more concerned about consistency between individuals. As simply rendered by historical evidence of kappa statistics, on studies that are entirely light microscope based, the kappa statistics give me pause.

DR. ADCOCK: Dr. Gilbertson.

DR. GILBERTSON: Color as rendered to the pathologist, clearly there's an issue. You didn't talk about it directly. I don't mean the color of the illumination but what is shown to pathologists, and we know it's a problem. Any comments about that? Is it a problem to have variation in color across different monitors, and if so, what's the issue?

DR. BALIS: It probably is a concern that needs to be characterized numerically by a perspective study, but there's both observational and anecdotal evidence to suggest it's not a significant concern. And I can give you two examples. One is, of course, the observation that there are a number of noteworthy pathologists that have significant color perception deficit in major forms of color blindness, and they're outstanding diagnosticians. So clearly a lot of the information intrinsic in the H&E is present in the luminance channel and not in the chromanus (ph.) channel, which is good news.

And then the other observation is that for any typical histology laboratory setting, batch reagents will often be used until they're exhausted. Sometimes between days, the hematoxylin and the eosin will drift slightly in tinctorial quality over time , and the ability of the human psychometric visual system to adapt to that is remarkable. Has that been rigorously characterized? No, but we all as practicing pathologists know an over-stained slide when we see one and usually will compensate for it. It's a very unusual slide in which we say this is so red or so blue I'm not going to deal with it. So I

think the qualitative answer is it's probably not a significant concern, but that doesn't reduce the need to perspectively characterize it.

DR. ADCOCK: Dr. Kulesza.

DR. KULESZA: I just have a comment more for the Panel than the speaker. I understand that the FDA and the Executive Secretary wanted to gather excellence on this Panel, and certainly I see very renown pathologists here, but maybe this is one example where you have some advantage, my relative inexperience, and that is when I took the boards, there were digitized slides. So already American Board of Pathology had moved on actually quite -- there was a component of glass slides, but the whole slide imaging was used to make diagnoses and test perspective pathologists.

DR. DAVEY: Can I make a comment? I'm a Trustee of the American Board of Pathology. So I'll disclose that. So we started doing this because it was very difficult to get enough slides especially for biopsies, and there were some problems with our initial imaging. I think it's gotten better, but still, we uniformly have complaints from the people taking it. Now, people complain about everything about the exams, and we expect it, but they always complain more about the virtual slides, and I don't know if that's a question. I would think the residents coming out of training would be getting more used to virtual slides and everything else. So I don't know if that's just a question of them taking the exam. I'd be curious for comments from our speaker about if that's a concern because a lot of the papers are

done by people who want to embrace this technique, and not everyone wants to embrace the technique. And I've participated in things from the CAP as well, programs with virtual slides, and some of them are very good and some of them, you can't tell what you're looking at because of the amount of material on the slide and how it's presented.

DR. BALIS: Well, without pointing to specific vendor platforms, one possible contributing factor to those complaints may have dealt with early whole slide imaging technology in which the acquisition platform had suboptimal alignments of the substage condenser so that the images had a soft quality because the spatial resolution was inadequate.

Again, in various types of whole slide imaging where the front end component is, in fact, a microscope, the quality of the image is predicated on how well the microscope is aligned. So some of those critiques or complaints may have been intrinsic just to the acquisition and not the technology.

DR. ADCOCK: Dr. Hewitt and then Dr. Sinard.

DR. HEWITT: As somebody who used it on the boards, I think it was the unfamiliarity with the platform initially that was one of the sources of problems. There's a learning curve issue.

So my question is you suggest an organ-based approach to the adoption of whole slide imaging. Is that the correct axis to think along considering, let's say, colon? I think colon cancer is pretty easy to diagnose.

It doesn't really require an enormous amount of magnification or -- complexity. But then as you wander into colitis, infectious colitis, in GVHD, you've introduced other diagnostic entities which may or may not be difficult. Same thing in skin. Melanocytic lesions, you point out, would be difficult. Other lesions might not be, but again, come back to something like infection agents. So is there another axis or means that one might be able to segment the practice of pathology because you always run the risk of having a glass slide on your microscope or your screen that was supposed to be one thing and turns out to be something else, and if it's something else and the technology is inadequate for that, you've got a problem. This comes to me from a rotation in pediatric pathology. You're not supposed to get pediatric tumors on your bench without having a suspicion that you did, and I got three in one month, and all of a sudden there were alarm bells throughout the hospital. There's a real problem. We're not doing our job right.

DR. BALIS: So to answer your question briefly, and then a longer answer, I think there are probably other ways that could be equally effective for organizing. The reason there's an attractiveness to the organ based approach is that if you look at other fields in which high reliability is absolutely paramount, again the aviation industry and avionics, if you look at the way for example a jet engine is designed, it's designed not to work in a range of speeds but at each finite type of torque, net thrust and performance. In other words, it's characterized and validated for that exact operating set of

parameters, and it works exactly well each and every time.

In the setting, for example, you mentioned like colonic carcinoma or colon. Then you have subtleties. So what about colonic with medullary variant which has a favorable prognosis or maybe it's involved by lymphoma. Again you could tease this apart and go from an orthogonal approach, not being organ based, but diagnostic area of concern, like dark cells that compromise your dynamic range or cells that have high need for verification of intracellular detail or structure. Again, these can be first principles based, and certainly that wouldn't be organ based. That would be based on the first principles of optics. Do you have the adequate dynamic range? Do you have the optical spatial resolution for this diagnostic entity? But at the end of the day then, you've gone to even a higher level of granularity, which is diagnostic entity based validation, which is now really every disease.

So I don't know that there's one best way to do it, but I think a good first principles common sense approach would be to make a list of those areas where there's concern and exclude them from routine use until they can be individually validated with the improving technology layers that are constantly being made available.

DR. SINARD: Hi. I've got a couple of comments that I'd like your feedback on because you've been in this field for a long time and have a lot more experience in this. I first want to step back a little, though, because

a lot of our discussion has been about how well the digital image can potentially capture the equivalent of the microscopic image.

But I think the real issue or question is the potential role of, you know, what is necessary to make the correct diagnosis, and we've all sort of alluded to it here. No one's actually said it, so I'll just come out and say it. But the quality of the image that you need to make the correct diagnosis varies tremendously based on what that diagnosis happens to be. And I think that's something that, you know, we were just sort of getting at a little bit here. There are some diagnoses that you can make with a magnifying glass that you don't even need a microscope for, and then others that no matter how good the microscope is, you're going to find pathologists who will want to sort of take it back to their microscope and look at it in their office because they just feel more comfortable in that environment trying to appreciate the --

You made the comment that the potential risk of inadequate image quality is that it could lead to a misdiagnosis. I'm wondering if you're aware of studies or anecdotal evidence or something in your use of this technology where the image quality was the cause of the misdiagnosis but the evaluator did not realize that the image quality was insufficient. Because clearly an assessment of the quality of the image is a key part of something that pathologists are doing all the time, and they're weighing that against what quality of imaging do I need to make the diagnosis, and is this a good

enough image. And the vast majority of the artifacts that we have to deal with and filter out don't come from the microscope or the imaging apparatus but rather from the sectioning or the staining. And so a key element of the evaluation process is what quality do I need for this particular diagnosis and that sort of thing. And all of that is fine because if we realize we don't have the quality we need, we send it back and we get more quality. And I guess the real risk would be instances where the quality that we need isn't there but we don't realize that it's not there.

DR. BALIS: So I'm not aware of any specific study that addresses the question the way you've posed it, but there is unpublished data which it may be moot from the perspective, its resolution perspective, but I carried out a hematopathology series with Dr. Carl Shellsburg (ph.) over 10 years ago. It may be germane to start publishing this data, and the pilot findings really tested that exact question. At that era, the highest resolution image, and this was not whole slide imaging. This was digital microscopy with a maximum resolution of 1280 by 1024 and 1600 by 1200, which at that time was an incredibly high resolution imager, and we looked at diagnoses for over 250 cases for which he had high confidence. So we looked at a confidence interval and then the accuracy, and we look at those cases where he had high confidence and yet he got the diagnosis wrong. And we noticed that it was a very small incidence, less than 2 percent. We're talking about small numbers of case. In all cases, they happened at resolutions less than 1024. So 640 by

480 or 800 by 600, which the reason I say it's moot, nobody uses a resolution like that in a contemporary setting.

So I suspect that that's no longer a problem in terms of not knowing that you're not looking at enough resolution. I think the greatest challenge is diagnostic acumen of two types. One, to be looking at the right field of view and not know the diagnosis or, two, not to be looking at the right field of view, perhaps because you're not comfortable with the digital medium and aren't as adept at moving rapidly through the subject matter as you would with the slide itself. And, of course, the third component is, well, potentially for those cases in which the dynamic range, the color range or the resolution are compromised, the information is just not there. But I think the majority of cases, it's more a question of simply sampling. You're simply not looking at the right field of view for a number of reasons.

I don't think the resolution is much a concern anymore at all because, in fact, if you look at the point spread function versus numerical aperture of a typical modern 2X or 4X objective, you know, which we do a lot of our time at low magnification, which renders with a Nyquist sampling frequency which is 2X as Dr. Descour indicated, you have to sample a 2X spatial resolution to get all the frequency information present. That's somewhere in the realm of about 2 to 3 to 4 thousand pixels squared, and that's getting within the realm now of what can be rendered on a combination of a two panel display.

So it is possible to render the full numerical information that is present in an optical image. So you could actually numerically show that with the appropriately designed digital panel projection system, you can show all the information that's invisible to the eye.

So again, I think the concerns are not that rendering but the quality of the acquisition that got you to the image and then on top of that, the dynamic range of the system, so that you're looking at the right brightness and contrast.

DR. ADCOCK: I think maybe one additional question. Dr. Foran.

DR. FORAN: So if it's not actually spatial resolution, do you think that field curvature error is something that gives rise to the drawbacks in digital microscopy?

DR. BALIS: So that's a very interesting question, and it actually begs the question of has any group designed a designed capture system, if you're talking about TDI based capture, based on cylindrical optics.

DR. FORAN: Yes.

DR. BALIS: Yes. Because if you were to do that, then you could completely get rid of the field curvature issue. I'm not aware of a cylindrical optic-based system. I do know, for example, in flow cytometry, the practice for normative focusing is to have one vertical and one horizontal cylindrical optic, to have iterative constraint of the focal spot.

I think, though, given how conservative modern optical design

is for objectives, with the internal field stop being significantly inside the capable angles of acquisition of the front objective, the perturbation in the modern -- objective is quite minimal. So even with the concern of spherical aberrations and the deviations from the planar performance, the proof is in the pudding. If you look at typical whole slide images, you don't have the scallop function where you have out of focus, in focus, out of focus and you just don't see it. Maybe it was present in earlier images. I think that's a minor concern with contemporary optics design.

DR. FORAN: But does it arise from the fact that you have a flat acquisition device?

DR. BALIS: I think the concern arises from the fact that, especially in architectures where you're using TDI acquisition, you're capturing strips which should be more amenable to cylindrical objects because your imager is a linear set of CCD or triple, you know, RGB, as opposed to a planar surface in which case you don't have to have a planar collection image, rendered image.

DR. ADCOCK: One final question. Dr. Gilbertson.

DR. GILBERTSON: Just a point which is: We're using the word resolution in various ways here. It's important. There's the resolution, to resolve two points which was talked about first, and then we talked about the pixel resolution on the screen, and they're very different. If I'm out of focus, it doesn't matter what my resolution on the screen is, I still get a blur.

DR. BALIS: Correct.

DR. GILBERTSON: So we have to be sure about that because it gets confusing quickly, confusing us. So they're separate things.

DR. BALIS: Yes.

DR. ADCOCK: Thank you, Dr. Balis.

DR. BALIS: Thank you.

DR. ADCOCK: We will now have a short 15- minute break, and I'd like everyone to resume at about 10:45.

(Off the record.)

(On the record.)

DR. ADCOCK: I'd like to introduce our next guest speaker, Dr. Michael Becich, Chairman and Professor of Biomedical Informatics and Pathology at the University of Pittsburgh School of Medicine.

Dr. Becich will be speaking on the digitization of glass microscope slides, hardware and software issues, features essential for interpretation of histology and cytology glass slides.

DR. BECICH: Okay. Thank you. I appreciate this opportunity to speak to this esteemed Panel.

You know, my background today and my title reflects my evolution as a practicing pathologist to focusing on the issues of biomedical informatics as they apply to this important domain space that I will sprinkle that experience into, but I have for my career trained as an anatomic

pathologist and was chief of a laboratory controlling both the microscopes and the digital infrastructure in a laboratory and its implementation.

So I thought what we could provide to the Panel is some lessons learned because we've been practically focused, and actually this very flat forehead that I have here, is from the experiences of banging my head against as many walls as I've had for so long, to try to move this whole process forward because at the end of the day, we're really responsible for improving the practice of medicine, and pathology is an important part of it, has been an extremely interesting place to articulate my role as both a physician and as a scientist.

So I hope today, although I changed the title of my talk a little bit, and I apologize for that, I can provide some lessons learned to assist in the thoughts about how the FDA and the Panelists should think about the issues that we've already learned, and the scar tissue is pretty deep. I won't take my coat off and show you.

So to start off with, I do need to state a number of disclosures. I have a national member organization that I'm involved with, the Association of Pathology Informatics and a national meeting called APIII that has focused in this area, has for the last 15 years, and we receive significant function from a number of corporate sponsors. I've outlined just some, not all of the companies from which I receive financial educational grants to run this CME meeting, that are in the digital imaging space, and I would outline them.

They're in your packet, and I disclosed them.

I also receive corporate sponsored research agreements from both Amgen and General Electric in this year currently and am part of a public start-up company, Omnyx, that is in the space which is a joint venture of the University of Pittsburgh Medical Center, the University of Pittsburgh Medical School, and General Electric, and I have several consultancies including one with GE Global Vision which I'll mention specifically. So those are my disclosed conflicts that the Committee knew about to start with, and I just wanted to make sure those were clear before I got started.

So to start off, I'd like to do a broad introduction to digital imaging and pathology, and there are numerous situations today where digital imaging in pathology has provided useful solutions to pathologists. It starts from the minute a specimen comes into gross rooms today. Actually today, gross room practices have been transformed almost completely by digital imaging. The documentation of resected specimens and requisitions has clearly been a home run for pathology departments in saving costs and providing a long-term archive of those very valuable pictures of gross specimens, medical legal cases that come in.

It's also played a role in the microscopic sign-out process, and today conference support is increasingly the educational conferences that abound in pathology practices, whether it's in formal case consultations, CPCs or actually nationally conference menus, have really benefited from digital

pathology.

And it's with that foundation that started early in the 1990s and really came to full being in the 1990s that provided the framework for what we are approaching today.

Then telepathology started to take hold with both digital and video cameras whereas digital imaging initially was all digital cameras and snapshot imaging, to then introduce various forms of telepathology, and I won't be talking a lot about that today, but I provide it in terms of completeness because the evolution to whole slide imaging is really a solution set that's evolved out of the inadequacies of snapshot digital imaging and dynamic telepathology, and I'll explain why that's the case.

So whole slide imaging includes, as we've heard, virtual microscopes and virtual slides. It's here today. It's the newest of these technologies. Some people make it, you know, draw some equivalency to what happened in radiology with PACS. For a number of years, there was teleradiology. It was off to the side and in a little corner on separate machinery, and then what happened in radiology practices, teleradiology disappeared because it all became digital. And I think pathology is in that evolution pathway.

It allows again, this technology, to have glass slides dealt with like digital files. They can be shared at multiple locations simultaneously and collaboratively, and with the drop in cost from network storage and

computing power, this is all really possible today. And so the huge implications for the practice of pathology is this is a real opportunity for patient safety, for pathology efficiency, and for adding a dimension to the way we diagnose that is not possible with analog use of microscopes.

So I'm not going to get into a lot of the detail, but I've always provided this as a backdrop of our activities at Pittsburgh. The line at the top that goes to 100 percent is gross pathology imaging. All the lines in the middle are different forms of telepathology that are employed at the University of Pittsburgh, but the one to pay attention to is the dark line, whole slide imaging, and if you look at absolute numbers of cases, whole slide imaging in our institution eclipsed all other forms of digital imaging in terms of numbers except for gross pathology imaging which has been removed from the slide, and you can see that it's growing very rapidly. So it's interesting to understand and dissect the trends of that process.

So why is whole slide imaging on the rise? You know, whole slide imaging as the basis for digital pathology today is really in a unique position to address several problems with conventional light microscopy, including the very time consuming and inefficiencies of snapshot imaging.

It also helps to share slides across distances very early, and in many practices, pathologists are distributed. They can no longer be anchored to one microscope in their office, and access to experts is also a growing problem. The ability to share cases with peers, the prioritization of case load,

the reallocation of cases across a network, the access to old images which is, you know, a dirty little secret of pathology. Pathologists rarely go back to the glass archives, not in the framework that they should, which should be pretty much every time you make a new diagnosis of cancer or you re-diagnose a malignancy or you talk about inflammatory lesions based on a slide before. The dirty little secret in the workflow today is we're not doing that as regularly as we should. This also is an issue that I think the whole slide imaging archive can help to address.

It also reduces and eliminates the issues with couriers, consults, filers. There's a lot of movement of glass that happens in pathology departments that is not really value added, and a lot of slides get lost and have to be re-cut.

So will whole slide imaging improve the quality of care? And these are my opinions.

Any new technology introduced to medical practice should improve the quality of care. We all know that, and as physicians, we're supposed to practice to the best of our ability, to the best of your patients, and avoid harming them, and those are the precepts that I think we have tried to move forward with.

If whole slide imaging is safe and effective, then many other applications will be possible, and I think the pathologist is an untapped resource to medical care delivery that this tool can open up additional

avenues to faithfully bring validated processes into patient care.

So let's take a look at validation studies today, and I will say I'm going to speak about three papers here and the authors, and the first authors of those three papers are in the room.

The first paper is by John Gilbertson, published in *BMC Clinical Pathology* in 2006, and this was the very first whole slide validation study of its sort. It was done at a very early stage. It used physical characteristics that are listed in the lower right. They're actually the same for all three of these studies except the first two studies are first generation type line scanner and the second and third studies are -- I mean the third study is a third generation line scanner. So I won't deal with that.

But this very early instance assigned three subject pathologists. They signed up cases. There were misdiagnoses in these cases as well as lesions missed, but it was concluded by the authors that whole slide did contain enough information for diagnostic use. There was consensus diagnoses, the original glass diagnoses which in this case was actually not a glass comparison to the digital whole slides but was simply using the original pathology report as the "gold standard" in this case. So there wasn't glass versus digital microscopy in this case. And it was concluded that, you know, whole slide imaging at this stage, again in 2004, 2005, and 2006 when this was published, was not as good as the microscope and that the focus really was a problem at this stage in its evolution.

And I really want to illustrate the exacting conclusions of this because this technology has really moved ahead from here, but in the framework of the things this paper taught us, there are several things we should consider in validation studies that we would design going forward.

And in the same spirit, there was a paper again in 2006, by John Ho, that was a quality assurance study, again another area of opportunity for whole slide imaging. In this case, there were 24 full GU cases that were looked at, and there were two independent reviews, glass microscope reviews and whole slide imaging reviews, and in this study they were found to be nearly equivalent, although there was a considerable issue here with whole slide imaging in that to review the cases, there was a significant higher amount of time required by whole slide imaging to look at this.

There were five clinically insignificant discrepancies, and in one case, the glass slide reviewer found a subtle carcinoma, prostate biopsy core, that was missed by both the original pathologist on the original pathology report and by the whole slide imaging review, and this paper actually includes an image of the missed carcinoma, it's 3 glands on a total of I believe 18 core images that were viewed in 1 part in 6. So it was a very subtle cancer.

But again, letting me drive a little bit into the complexity of the reviews, and also touching on the issues that interobserver and intraobserver variation would be a key thing in terms of thinking about how we validate diagnostics since there's a lack of gold standard and an inherent error rate in

glass microscopy which this study helped demonstrate.

The final study is by Jeff Fine in 2008 that looked at whole slide imaging and immunohistochemistry. In this case, 30 difficult prostate biopsy foci were studied. There were five subject pathologists and one outside expert. There was a review of whole slide stains. They used the kappa statistic that has been mentioned a number of times in the previous speakers for interobserver agreement, and the interpretations were that the pathologists saw the same thing but disagreed on occasion on how to interpret it, again outlining the inherent flaw of interobserver and also intraobserver variation that should be taken into account when we look at it.

So what are the lessons learned from these early validation studies?

So there are issues that relate to the image capture or the hardware lessons learned. Results from these early studies were first generation imaging robots using line scanning. The images were good enough, a focus group consensus, from the immunohistochemical studies for interpretation. They were able to do hard cases, and one of the real pearls of this is pathologists could figure out when it was time by looking at a digital image to defer to a glass slide.

So I want to start to introduce the concept of whole slide imaging does not replace the glass slide. The glass slide is still there, and just like when a pathologist looks at a glass slide and can decide that they want to

get a re-cut or they want to get it re-stained, whole slide imaging doesn't mean the glass slide goes away. And the pathologists are the most exceptional filters of noise. I've never seen a discipline that could look at noise and know when it's there and know when to default. In a whole slide imaging environment, I think we should promote the concept that the pathologist remains in control and can go back to the glass slide and can do all the things they do today, but what are we going to bring to the table that will be additional benefits to them?

So in terms of the digital pathology workflow of software lessons learned, these are a lot of observations by Dr. Gilbertson in his 2006 paper that were also echoed by Dr. Fine in his 2008 paper. There's a lack of a digital slide box, meaning looking at the slides, in the initial software that was screened, you couldn't really see the slides. One of the things pathologists do is they just eyeball the slides, and you've got to have that capability to look at just a snapshot or the actual virtual slide of the slide without the magnification.

You can't preview a case prior to opening the slides? Well, you know what? Pathologists are smart. When I sign out with a resident, I know the two slides I have to go to just by looking at them on the tray. So we've got to have that tray, and you've got to be able to preview, and you've got to be able to do these things.

So whatever the software does, it can't steal away from all of

the information a pathologist gathers during sign-outs. So the design of the software has to allow for that.

In the 2008 study, there were issues with low and medium power magnification. There's wasn't enough monitor real estate in the studies. They didn't control for monitors. As you saw in the previous three studies, there was absolutely no control for the monitor, the monitor resolution, nor for the light microscopes in these studies, and only two of the studies actually used light microscopes. The other one was just back to paper. We want to make sure we control for those in any validation studies.

So the lessons again learned from a design perspective is we need diagnostic standards to design better study conditions for head-to-head comparison of light and lack, and the problem is in these current studies, those things weren't considered, but they really shed incredible light on what we need to do in the work we are proposing today.

We must consider intraobserver discrepancies and identify those causes. Many times these discrepancies, even in the studies that we have performed to date, indicate it's not the image but are quality related on things outside of the control. So we have to be very careful on the design of our studies.

Subject pathologists sign out cases they're accustomed to signing out. When we do a validation study, we should understand that general pathologists sign out general cases, and specialty pathologists sign

out specialty cases, and that workflow is not equivalent nor is the focus nor is the toolkit. So this has got to be a consideration when proposing retrospective validation studies.

Subject pathologists see slides scanned from histo labs that they are accustomed to seeing. Different histo labs have different cutting and staining protocols and things that vary, the usual workflow of pathologists can lead to artificial discrepancies. That's going to be very difficult to control, too.

And again, early off, no considerations were paid to the software used to present the whole slide. You got a whole slide image. It was just an environment for -- image. You know, pathology sign-out is not that way. So we have to look at that issue as well. It will probably have a profound impact not only in the speed but the adaptability of the diagnostic interaction.

So physical characteristics which a lot of the folks have outlined already, and I won't spend much time in here. You've got to look at the device type and the resolution, the objectives used, the monitors used, the monitor resolution and, you know, if we're going to really do head-to-head comparisons with light microscopes, which probably isn't the way to go, we're going to have a whole bunch of features that are not controlled in the practice of pathology today, meaning all those things we've heard by the previous two speakers.

And, you know, how do you then do a comparison of whole

slide imaging with light microscopy when there's no software environment that goes with light microscopy imaging?

So these are all issues that will really be important to design carefully.

And, you know, once we prove that workflow software enhances the pathologist's effectiveness and patient safety, can we avoid using that in our environment? So that's a downstream effect of whatever decisions this Panel makes.

So, you know, again this is a very important driving concept. It stung a little when the introduction to this Panel said that we're looking at whole slide imaging to replace the glass slide or light microscopy. That is not what I personally advocate. The pathologist is in the driver's seat. The pathologist will be relied upon and determine when glass slides are needed. This is not going to eliminate the glass slide. This is a glass plus environment. Whole slide imaging will be an adjunct to glass slides in that case, and in instances when the glass slide is not available, and the pathologist needs more to make a decision, then he should defer just like they do in general practice today. It happens very commonly in frozen sections. It happens in biopsies. It happens in cytology. It happens in whole arc case diagnostics, and it may be that you give a diagnosis that's deferred because you can't get the other information. Whole slide imaging won't change that. And I think again pathologists are very, very effective at making those decisions today.

So in the scope of discussion is obviously the digitation process of glass to digital image, there's also the workflow tools, and I want to extend that, you know, it's going to be very hard to do head-to-head comparisons of just the imaging when the software may be bringing additional features that you cannot compare in the microscope environment. And I do think out of scope is there's much more variability and this is, you know, coming from 25 years of pathology practice in the histology lab with what comes to you on the tray than there is with the imaging component of this. So I just want to highlight that.

So what are the innovations with the software? Again, these are my opinions. The retrieval of the previous biopsy images. It's a major safety issue, and standard of care that should be a major goal of surgical pathology that this technique will facilitate. It's been one of my personal drivers in pushing the discipline in this direction.

It needs to help facilitate informal consultation. My number of consults was extremely dependent on the time I had in the day. That's not good for your mom, my sister, your dad, my child when a pathologist is under those time pressures and can't share a case instantly. And when you have to walk across a campus or go up a floor or even try to schedule another busy pathologist in the way, that's not safe practice.

I think that this is one of the mainstays of high quality surgical pathology, to completely facilitate formal consultation and this technology

will do that.

Distribution of routine work is also important. You know, we want the generalist doing general pathology. We want the specialist doing specialist pathology. They can really help one another if we can distribute the workflow.

And then finally, real quality control, one of the most uncompensated but important parts of surgical pathology practice that in my opinion is not regulated enough is true quality assurance within the practice of this discipline. Digital imaging could provide a very cost effective way for distributing, and with the early validation studies already showing the best results with whole slide imaging, I think this is really an important part to pay attention to because it will promote safer practice.

So the final points, I agree with Dr. Balis. Instead of saying all of hematopathology and all of cytology, I think that there are cytology type preparations that are inherently 3-dimensional that should not be considered in whole slide imaging because of that until we have Z-axis scanning. So that doesn't mean all of hematopathology should be discarded. There's certainly lymph node sections that are based on plain or images in paraffin that can certainly be used in whole slide imaging, but needle aspirates, any 3-dimensional based specimens, hem smears which are more 3-dimensional than most, although there's digital imaging applications used in automated hem screening, we've got to be careful about what we say, and I actually think

the approach should be what are the physical limits of the imaging itself as the yardstick instead of any particular organ type or organ line.

We did some studies back about a decade ago with engineers at the Pittsburgh Super Computing Center to say how many images would represent the entire practice of pathology. Just have one ideal image. It was 600,000 images, and it was probably an underestimate because it didn't include special stains, special diagnostics. If we go into the complexity of trying to go down all of the path of what we see, what a pathologist sees and makes in diagnosis, it will be far too complex to think about how you would go down each of those paths to make this happen.

So again, I advocate for leaving the power in the hands of the pathologists, letting them know when they can use digital glass and when they need real glass, and that the two work together.

So oil immersion is a clear carve-out, and we covered this, so they're not capable of doing this, but, you know, you'd understand that this is less than one-tenth of one percent of the entire practice of anatomic pathology. So it's a very, very small carve-out.

So final points, we don't have reference standards in light microscopy. They don't exist today, and there's tremendous practice variability. Anecdotally when I took over as Chairman of the Pathology Department at Shady Side Hospital, that was converting from a private practice community to a very acute care cancer hospital, I went down

to the pathology laboratories and found one wide-field microscope, that had two heads on it, and it was the only one that was possible of Kohler illumination in the entire department, and I freaked out. I insisted until we replaced all the microscopes in the laboratory, that no cancer diagnoses, no new diagnoses of cancer be done except on my scope in the laboratory, on that two-headed scope. The inability to regulate what happens for people. I mean actually medical students in my medical students class at medical school had better microscopes than this community hospital I was now chief at.

So we've got to realize that standards of care delivery are not reinforced and our colleagues probably at the College of American Pathologists should step up in this regard.

There is commonality in some, you know, both of these methods today use objective tools. We're not really talking about changing the way that we acquire images from glass. We're just changing the way we distribute them.

Now, that's an oversimplification. We know that it is. But, you know, every one of these robots you open up today has got an objective lens, use light properties. It uses all of the physics we talked about. So it's a good foundation for standardizing optics in all of pathology.

So, in conclusion, we should firmly focus on the ability to diagnose and resolve histologic structure irrespective of the method of the

images presented, and I think that's the analogy from what happened in radiology.

So my final points are, are there things that we can evolve as imaging reference standards? I personally favor a diagnostic yardstick for validating whole slide imaging. Since a reference gold standard doesn't exist, maybe we take and we put certain visual -- that are used across the diagnostic disciplines, and say these are the 20 or whatever number of things that we need to be able to see with any imaging system to validate it as an adjunct to diagnosis, and not replacing the glass slide or changing surgical pathology practice and the FDA in their notes suggested cancer invasion of the basement membrane, cytoplasmic vacuoles, cross striations, micro-calcifications. There are other things like gram stain organisms, immunohistochemistry, chromogens, eosinophilic and basophilic, nucleoli, mitotic figures. These are the things pathologists see and assimilate in making a diagnosis. Now, this is not a comprehensive list, but I think it would be very worthwhile to actually go to the end game, the image produced at the end and say can we detect these things that go into making diagnoses.

It also might include some hybrid features that are based on imaging like color and refractive properties, but those will be alien to diagnostic pathologists.

I think the practical approach to really real in the practicing pathologists in the community, which is 85 percent of all the diagnoses

rendered, is to use the things they use at their microscopy every day and to do the hard work of compiling those.

I was introduced to this concept of phantoms as a result of the background work here, and these are interesting and also a concept worthy of exploring. Looking ahead at some of their slides, I think we're going to hear about some other potential ways using artificial, meaning non-human tissue based phantoms. I think this is a goal, but to vet one of those standards is a very long process.

On the other hand, I think it's a grand challenge, and there are communities on biomedical informatics. We could organize them in pathology informatics because we have all the right partners at the table to help think about phantoms that would be applied to an instrument, but it's going to take time to do this.

On the other hand, I am -- almost all the pathologists in this room are members of one of these organizations or many of these organizations, and I think the thought processes are coming very close together to doing this, and I think we should start a dialogue on that including APIII, API, and the Digital Pathology Association, but also we formed a great partnership with the Society for Imaging Informatics and Medicine which is a radiology imaging group. They now co-present scientifically with us. The College of American Pathologists and Futurescape have been absolutely instrumental in moving us along this path. So all of the aligned entities and

even the USCAP, the Academic Surgical Pathology Association, are all really focused in trying to move things forward. So I think we could draw together the critical mass from the pathology community and the radiology community and do this in a fashion that would be efficient.

So again, my conclusions are whole slide imaging does not eliminate the glass. The pathologists are very effective in knowing when they need more information, and we need to continue to rely on them. This does not take that power out of their hand. And again, it doesn't in anyway replace the pathologist's critical judgment. It's an adjunct and an enhancement to the way they make diagnostics.

Again, I've done some carve-outs particularly in 3D cytology preparations, and those require oil immersion diagnostics today, but I think again the concept that whole slide imaging will mature and may surmount those barriers is a possibility, but not today.

And there are other areas of diagnostics that should be excluded, but again the whole slide imaging pathologist will know when to go to glass. They do it already. They've done it in our very early studies.

So I think again the compelling reason to move forward is to improve the quality of care delivered, improve the patient safety that we can evolve for our patients, and aid in the evolution of surgical pathology, and these are all my opinions. We've been pushing in this direction because we very firmly believe digital pathology tools to this workflow has significant

advantages over light microscopy and will meet the goals that the FDA is a vanguard for.

There are real challenges in essentially starting with a non-standardized, non-gold standard practice and to try to raise the bar, but I think, you know, we've got the right people around the table to do this. I like the idea of these diagnostic yardsticks, the evolution for diagnostic phantoms, and I welcome the input and comments by this esteemed Panel and presenters and public.

There are a number of the presentations that fed into my talk today. The three studies that I outlined are bolded in red, Dr. Gilbertson, Dr. Ho, and Dr. Fine, who are in the audience, and I want to thank folks. If they have any questions, feel free to e-mail me. I'd be happy to distribute this PowerPoint to anyone who would like the actual presentation. So thank you very much.

DR. ADCOCK: I'd like to thank you, Dr. Becich, for your presentation.

And at this time we would entertain some questions from the Panel members. Dr. Gilbertson.

DR. GILBERTSON: You showed a number of studies. I know them all very well, and I have kind of a question because something bothers me. We've shown that a group of pathologists looking at these digital slides essentially created by a robot with one focal plane often not in focus, with a

strange monitor, monitors not being controlled, a whole bunch of stuff that really is kind of weird, okay. At 20X magnification, we're able quite realistically to give the same diagnoses as pathologists who basically signed out the case, and I mean with the glass slide, you know, with stuff on the line, all right. I mean people's bodies on the line. Very impressive, but it's also kind of weird. Are we saying that pathologists don't need 40X? Are we saying the pathologists don't need to be able to, you know, focus on none of our specimens? I'm not quite sure that's true, but I don't understand why we're not picking that up --

DR. BECICH: Yeah, again --

DR. GILBERTSON: -- in terms of study design. Why aren't we picking that up?

DR. BECICH: As much as the 20X studies that you, Jeff, and John have done indicate a bright light at the end of the tunnel, I do think that there's a lot of other discovery that needs to happen. I actually think it's kind of amazing and very hopeful that we're actually able to conclude that with this low end technology, with the lower end objectives and without optimal viewing conditions and without the right real estate on the monitor, but I owe that to just one thing. That's the incredible ability of a diagnostic pathologist to know when they need more, and again, that's what this environment's going to provide.

I think the -- studies, the inoperability disagreement, the focus

that was missed on those studies probably will all have been avoided by 40X with better software. And so my conclusions are, and why we're pressing forward with this today, and I think the FDA has wisely got us sitting around the table to help evolve a way to do this is that the time is right.

DR. GILBERTSON: Let me ask my question a different way then. If I took all the pathologists at Mass General, took out their 40X objective lenses some night, okay, would the quality of the diagnosis, the pathology done at Mass General go down?

DR. BECICH: As long as your senior pathologist, and these are studies that, you know, we've done at Pittsburgh, senior pathologists rarely use the 40X objective.

DR. GILBERTSON: They all have them.

DR. BECICH: High power mine, diagnostic powers are 95 percent of the time at 2 and 10X. Only 5 to 10 percent of the time at 20 and 40. You take a junior pathologist, and the curve flips. Junior pathologists need 20 and 40X almost all the time.

DR. GILBERTSON: That's my question really is just, it's obviously a kind of a strange question, the questions is why aren't our studies picking that up, and I'm just wondering.

DR. BECICH: Because I think if you have a high powered pathologist, they need low powered objectives, and if you have a low powered pathologist, they need high powered objectives.

DR. GILBERTSON: One more important question on this. You didn't say one important thing about these studies.

DR. BECICH: Okay.

DR. GILBERTSON: That I think in all of them, at least some of them, the robot in a significant amount of time, 6, 7, 10 percent, failed to image all the tissue on the slide. It failed to find the tissue, and this is an important thing. I think they can do better, okay, but it's a serious issue to remember because that's the one time the pathologist will not be able to use his superior brain to say I need more because he never saw more. It's not there.

DR. BECICH: Agreed.

DR. HEWITT: Mine's a small follow up. Have any studies looked at non-neoplastic processes effectively? I mean Dr. Gilbertson's comments about why we need, why is 20X working, and I sit back and I go, you know, I agree with you that if you're an experienced pathologist, you need a low power lens and you can do your job very effectively, but you wander into nephropathology, transplant pathology, you need a 40X to look at, I assume metravacuolization (ph.) for cyclosporine toxicity and things of that nature. And so it's interesting to me that, and I'm from the NCI, everything seems to be focused on oncology where some of the really hard questions may be well outside of oncology.

DR. BECICH: I agree, and infectious disease, this new

theranostic era that we're entering into, where we're going to be looking at disease interval and making judgments on it, is another pioneering area. So I agree, there are unknowns, and the studies need to include a broad array of diagnostic entities and not just oncology. And I'm sorry we're so prejudiced about cancer at the University of Pittsburgh. Not really.

DR. ZHOU: So I have a few questions about your three reported studies here. The first one will be easy. I think the kappa you use is not correct because you have three -- correct when you should use a weighted kappa. So I don't know whether that will change the results or not.

DR. BECICH: Say that again. I'm sorry.

DR. ZHOU: Used a weighted kappa because that's not binary data. You have --

DR. BECICH: I wasn't actually a co-author on these studies. Dr. Fine, can you make a comment on that or is that -- is that all right with the Panel, if I have the author actually answer it. Come on up, Jeff.

DR. FINE: My background is not statistics, but as I understand it, the kappa statistic that we used was essentially an attempt to quantify what the probability of or quantitate the agreement while correcting for the probability of just chance agreement, very simplistically.

DR. ZHOU: Yeah, I mean the original kappa is for the binary.

DR. FINE: Yeah.

DR. ZHOU: Is the answer yes or no? But if you have a --

category, here is substantial moderate -- and then one and three and two and three are different. So maybe the weighted kappa will be better way to go.

DR. FINE: Perhaps. And again, we were dealing with basically comparing two people at a time, and we relied on literature, other surgical pathology studies to try to figure out what would constitute good or great agreement, good agreement, or marginal agreement, going by what was already in the pathology literature trying to determine what would be I guess good cutoffs for kappa values.

DR. ZHOU: So my second question is your study by Ho, there are two independent reviewer, the glass and WSI. Are the same reader do both reading or different reader do different readings?

DR. BECICH: I believe those were different readers, weren't they, John? Different readers.

DR. ZHOU: Yeah, so that raise the issue about how do you separate, the agreement is due to reader effect or due to actually assistant effect, the image itself effect.

DR. BECICH: I agree, and that's why I bring these up. These are not perfect studies by any yardstick, but they represent the complexity of issues that we need to do in a very well defined validation study. So again, we're not holding forward either the findings or the framework but to use them as a platform for the right design going forward.

DR. ZHOU: Yeah, last question I actually open up for everybody

here is about how do you -- what you going to do if you don't have a gold standard which is relevant to here if you want to compare the glass and the WSI? Actually, in statistic, we do have some methods available which allow us to -- to estimate -- curve if you have a -- reader read the same cases, let's say at least five readers read the same cases under one condition. I don't know that condition reasonable hematopathology or not. So a condition we have to make is that condition on true disease status, the reading from different reader are independent. So that's something we have to make. If that assumption not be true, in some study, actually we able to do estimates -- without gold standard. So I want to -- assumption --

DR. BECICH: I think we have to -- again, I'll make a comment, and then if anyone else would like to comment, that would be fine. I think the first thing we have to do is we can't control for all of the physical parameters of light. I mean the complexities we see are what indicate that we're just not going to be able to make equivalent what happens in a non-standardized pathology laboratory with a pathologist reading a microscope on multiple microscopes to whole slide imaging. So, you know, to me, I'm going to focus on the actual quality of the image at the end of the game and the ability to diagnose discrete objects within it because that's really what we're, you know, what we probably have to focus on, the fact that all the rest of what's in the middle is a black box, and again there are precedents for this in the automated cytology industry and in radiology that may lay some track

along this path. But that still is a complex problem to control, and we're going to need statistical input in both the design and in the analysis to get it right.

DR. MELLO-THOMS: Dr. Becich, you said that the different histology labs cut and stain the slides differently. So how do you see that there's an impediment to standardizing color for WSI, namely as we've been told that there are variations currently between 10 and 15 percent amongst the different pathology labs? Would that be a problem in your opinion? And you think that in order to solve that problem, you would need a digital correction to the images, or you would need the more standardized way of cutting and coloring slides?

DR. BECICH: Well, that's a complex issue. If I were to be asked, which I guess I am, the way I would try to do it is to try to as tightly make sure, whatever comparison we proposed, is as close to optimum as we can get it, knowing there isn't a gold standard but knowing that we don't want it too blue, we don't want it too red, we want it just right.

I think the most important thing that all of this has taught us that, okay, so what if my lab always produced red slides and now that's what I'm used to, and somebody else's lab produced bluer slides and that's what they're used to. I think the pre-study training period on the subjects is going to be a very important part of this. It's got to eliminate unfamiliarity with whole slide imaging itself. It's got to eliminate the gray zone of lab to lab variability, and it's also got to address who are the subjects because we can't

get specialists who just do whole slide imaging, and we can't get generalists who have never seen a whole slide and mix them up and hope that our results at the end are going to be very worthwhile.

So I'm not going to directly answer your question about what we do about color. What we're going to have to do is try to tightly control it in an environment where everybody has practice and is efficient with the use of the technology employed before any measurements are made.

If I had to kind of criticize all the work we've done at Pittsburgh to date, we're still neophytes with this technology. We're learning as we go. We're bold to put our work out in the literature to essentially create parameters about what should come next, but to get to where we are now, I think there's a lot of scar tissue and a lot of lessons learned that can be used to employ a much better study than it's ever been there, and when we go to validate for, you know, 3-dimensional subjects, with Z image, we're going to know even more then, and when we go to oil immersion and 100X, which will happen down the road, we'll know even more then, and what we should do is continue to push the bar forward and perhaps have an impact on the conventional practice of light microscopy because this is not going to happen overnight. There are obviously issues that we will empower the discipline of pathology and the FDA to do that will -- a perfect practice, and that's our job at the end of the day.

So just to commit to doing the best we can do is the way to go

knowing we're not going to get to perfect.

DR. ADCOCK: Dr. Davey.

DR. DAVEY: Yeah, just a comment. I agree with you in terms of primary diagnosis of cytology images probably shouldn't be considered. However, I just wanted to make sure that you wouldn't exclude things like immediate evaluation techniques and things like consultation and archiving, and I'm just curious about your experience with that because I think a lot of cytology, because a lot of cytology laboratories are using either telecytology, but they could use imaging for immediate, sort of like, you know, analogous to frozen sections.

DR. BECICH: I'm so glad you asked this question because I was hoping somebody would. I had to cut my slides to make sure I did 40 minutes.

One of the studies I didn't talk about because it wasn't whole slide imaging based was the excellent work our neuropathology group has done using even more primitive resolutions and even more primitive technology than whole slide provides. And we do not have a 5-year experience, and we do 25 percent of all our neuropathology frozen sections. The key is, it's all done by touch prep and, you know, touch prep is inherently 3-dimensional like the things we want to exclude. So again, I'm nervous because unless we cast a broad net, something at the fringe might make its way through. So probably my initial recommendation would say anything that's 3-dimensional shouldn't be used.

On the other hand, we've got this problem in that today people are doing telediagnostics on cytology and neuropath as a regular part of practice, and we could really impact, we could shut that down, and I think it's very important for those disciplines that we not. So we've got to be careful.

So as much as we think there's a black line and it sounds like there's agreement that 3-dimensional specimens should be eliminated, you know, what about thin prep?

DR. DAVEY: I think for final diagnosis, it's different, but I think for preliminarily, yeah, I would agree with you. You have the option of deferring all the time like you would have in a frozen section.

DR. BECICH: But again, my concept, the power of the pathologist, give them more tools and let them decide when digital is not right, and if you do that, and you add the value of whole slide, we have better pathologists at the end of the day. I'm convinced of that. But that's my opinion, of course.

DR. ADCOCK: Thank you, Dr. Becich.

DR. BECICH: Thank you.

DR. ADCOCK: I'd like to invite our next speaker. Our next speaker is from the FDA, Dr. Kevin Lorick, a Scientific Reviewer in the Office of In Vitro Diagnostic Device Evaluation and Safety, the Division of Immunology and Hematology Devices. He will speak to us on the basics of digital imaging.

DR. LORICK: Hello, and thank you all for being here today,

especially our distinguished Panel members. I'm going to speak to you today about the basics of digital imaging, in particular, digital image capture and display.

In the context of today's discussion, whole slide imaging is the creation of digital images from glass slides. The creation of these images depends upon computer technologies whose purpose is to convey to the digital image the same information that's contained on that glass slide.

Digital imaging is the collection of images on a pixel-based collector. Now, that pixel-based collector is typically a charge couple device or CCD, or a complementary metal oxide semiconductor, CMOS, with or without a photo multiplier tube added to increase the sensitivity of low light signals.

Now, pixel simply put is an element of a picture, picture element. This can be very typically the dots that you see on a printed picture, the display that you find on your monitor, or in the case of this discussion, it's the micro detector of units that are found on a sensor.

Pixels themselves have spatial area which is essentially a shape and size. The size is typically described in the pixels per inch that are found on that sensor. They have issues of color. These are referred to usually in bits of information. So one bit colors the simplest form, pure black and white. You can get 24-bit color, which is 16.8 million colors. It's also referred to as true color. And each of these particular sensors have different sensitivity to

photons. This is particularly important at higher magnifications and when looking at wavelengths of light particularly in the reds.

Now, in order to digitize a microscope specimen without losing any details available to the human eye, the dimension of the detector element, that is the CCD pixels, should be no later than one-half the corresponding limiting resolution distance. This is the 2 pixel requirement that we've heard about a couple times already. This particularly applies to plane or flat images.

Now, as Dr. Descour mentioned earlier, the limiting resolution for a particular lens is determined by its numerical aperture. For example, a 20X lens requires a numerical aperture no greater than 0.5 NA. The limiting resolution of this lens is approximately 0.8 microns; .8 microns will fit into an inch 31,250 times. If you require a 2 pixel grid for that, you need 62,500 pixels per inch.

For a comparison, a high end 21.1 megapixel consumer camera contains less than 4,000 pixels per inch or about 1/16th of the information required in one dimension.

Now, that said, this doesn't necessarily apply to -- it applies to the information that's available to the eye. It doesn't necessarily apply to the information that the eye can discern.

The pixel sensor at its base is actually a monochromatic device. It doesn't actually capture color. So there has to be an additional technology

added in order to capture that color information and distinguish the independent colors coming into each of the pixels.

The most common method used is the bare filter, and that's shown on the left. The arrangement of the bare filter is actually shown in both of these cases. On the left is the more typical square pattern. On the right is something found in more modern sensors, this honeycomb pattern with offset 45 degree pixels.

In the bare sensor, each photo site or actually in this case, each pixel or the capture pixel, the arrays consist of a single light sensor, either CMOS or CCD, that as a result of filtration is exposed to only one of the three primary colors, green, red, or blue, and you'll notice that there are two greens for each red or blue. This is because of the sensitivity of the human eye to green light.

Now, one of the practical issues that comes up when we talk about the resolution capture of two pixels, you'll notice that in this particular very common arrangement, the 2 pixel by 2 pixel arrangement is the smallest unit of that particular sensor that can capture all the colors of light.

One thing about the bare filter is that it is, we were talking about before, damaging in terms of its ability to capture all of the light. The bare filter will actually filter out about 70 percent of the light that's coming into a CCD sensor. So other methods have been developed in order to capture color, and these include the use of the rotating filter, which

sequentially filters out red, green, or blue light before it hits the monochromatic pixel sensor, the beam splitter that actually directs the three different wavelengths of light to three different CCD sensors -- this is a much more expensive way of doing things -- and then the additional ways, the scanning mask or filter that uses pretty much the similar methodology, the rotating filter, in that it captures each color of light sequentially.

There are other technologies that are available. Somebody earlier mentioned the silicon chip that can or vertical pixel that can capture all the colors of light and discern them within the single chip.

So how can the CCD sensor detect luminosity, and that's essentially what it's doing. It's detecting the luminosity of light hitting at each pixel after filtration, and it's recording the photons that are approaching it.

So different wavelengths of light appear identical in luminance when converted from the color HSV cone here, which is for hues saturation and brightness value and converts it essentially to the brightness values seen on the right side. The black and white can be then recorded as 10 logs of intensity, and each pixel may record the luminosity of each color in terms of bits. Basically the color becomes an instruction to software that's attached to the sensor. Most people think of binary language in computers, and that's exactly what a bit is. It's one digit in the binary language.

So essentially what a bit is is a binary digit. So the simplest 1 bit method has 0 or 1 value for the single binary digit. A 2 bit instruction has 2

values. In each case, it's always 0 or 1. A 3 bit has 8 values. And old black and white had 16. One of the standards people used to deal with, the 8 bit, has 256 possible colors or 256 possible values. Each of those values represents a color. The current monitor that you probably have on your computer is 24 bit, and it will have 16 million possible colors.

Now, the tradeoff in recording a higher bit depth when it comes to your pixels is obviously storage capacity of the image.

There are other issues with image acquisition that have to be considered. First and foremost, are we dealing with a single image or a composite image? Most of the time in the context of whole slide imaging, we're going to be dealing with a composite image that's going to be a Mosaic that's stitched together. Is the image there for publication or is it for analysis and diagnosis? What magnification is the image recorded at? What format are we recording? File types, there are a number of different file types available for imaging, including TIFF files, BMPs, PNG, JPEGs, etc. The image can be color or black and white. For the purposes of the sensors, black and white has some advantages of clarity and resolution and more information, but because of what we're talking about today in whole slide imaging, there's always going to be H&E stain. So we really want to capture a color image.

Also you want to know if you're dealing with a single image or single specimen or series of specimens. What type of resolution, and this is an important one, is required to obtain the information needed to make a

proper diagnosis. What image processing is going to be required in order to convert the pixel information to a computer file and then back to a display image? What types of compression are used? And then also what are the basic camera settings? We have the camera in focus. Have we calibrated it properly? So some people touched upon this.

Again one thing to look at is the camera. Essentially, is it in focus? Is the resolution correct? Again, you can look at a test pattern like the one on the right. You look at white correction or white balance. Is what we are recording as a white color on the image the true white? And is what we're recording as the black image the true black? Can we correct for background? The biggest image when it comes to imaging in general is going to be elimination.

And this is just a quick slide to point out that changes in calibration can affect what we see in the mid-tones of a particular digital image. So if you go too far to the right, or too far to the left, you're going to get information that's not a true color. If you don't go far enough, you're going to eliminate some of the subtleties of the image.

When we start dealing with image capture, we also have to deal with storage, and because of the size of the images that we're talking about, this is becoming less and less of an issue, but it's still a critical point. A typical 1.4 megapixel image, which is the old digital camera you might have, would require 1.4 megabytes of storage space if you were storing in 8-bit color. The

example I gave you before of the 21.1 megapixel camera true color image requires 63.3 megabytes for a single image. To achieve the ideal resolution of that 20X objective lens that we've talked about before, you potentially would need greater than 16 gigabytes. Other alternatives that may be required could be time consuming and/or processor intensive or software intensive.

So part of this comes into capturing your field of view. What the camera images in the microscope is not the whole picture. All right. So in this particular example, the round image on the right is what you would see in the ocular of the microscope. The square box or the rectangle in the middle is what the camera is actually going to see in a typical microscope setup. Because the camera is a physical attachment to the microscope in this particular case, you will find that it will tilt, gravity will win out, and it will be off center in some cases. The result of this is that despite the best Kohler illumination, you're going to get some degree of shading of the image.

The other thing about the digital image itself is that if you want to capture a larger of the specimen than the field of view as permitted by the camera, you're going to need to do some type of collective or repetitive image collection. Now, one way to do this is by simple point mapping. You look at one field of view in the microscope, you take the picture, you move to the next one, you take a picture, you move to the next one, you take a picture. This could be slow, and it gets into issues of spatial resolution.

The second way to do it would be to go through line mapping

either using a scanner or potentially linear ray of lenses. It's a little bit quicker and has better spatial resolution.

Or, you can try global imaging where you take an array of microscope lenses and image all at once.

One of the questions when you go to the faster imaging is are you increasing the signal to noise ratio. So that's also one consideration that has to be in place.

This is just an example of an image display, and the idea that just because you use a digital zoom on the image doesn't necessarily mean you're creating any more information, but if you don't use enough of a pixel depth, you can, in fact, lose that information. This is not very bright on the screen, but you'll notice on the top left a low magnification of a tissue microarray essentially taken at 232 pixel squared.

If you blow that up to a much higher level, you'll notice you lose focus, and it's down on the bottom left. If you then increase to a 20X magnification, and you use a 464 x 464 pixel capture, and now you blow that up, you get a little bit more information, but it's still out of focus. Then you can go all the way up to 3,090 x 3,090. That's getting closer to that 18 megapixel image capture, and you begin to see essentially all of the detail that you need.

But when you get into those larger images, one thing will happen. It gets into another tradeoff. The smaller image can show you the

information you need using the smaller pixel depth, the 232 x 232, and it doesn't need to be compressed in any way. When you start using the larger images, and you start having to sew a number of images together, you get into problems of storage, and that gets into the image formats that you're using.

So the simplest, easiest imaging format is uncompressed, meaning that each pixel records with full color depth, but the file size can get very large. You can use forms of lossless compression where color values of the pixels are unaltered. And then you can use lossy compression, and this usually involves some form of Fourier transform. Lossy compression will result in the reduction in spatial detail and reduction in number of colors recorded. That may be done by removal of colors that are rare and not present in the image, or in a severe case, you get decreased bit depth of your color.

Now, as an example, this is my best friend. Give me a hand here. This is a microscopic example of image compression. The photo here is a 6 megapixel image, and it's been compressed slightly down to 678 kilobytes using JPEG format. Now, here the image has been compressed slightly, and possibly the people a little bit closer can see it better. I don't know if the people in the back of the room will be able to appreciate it, but there's a slight loss of focus.

Now, if you go and you blow that image up now, you start to

see a loss of information. In particular, again if you're close and you have a little bit better, this gets into some of the issues that we're talking about today as a matter of fact. This is a much darker display than the one that I'm looking at right now, but on the left, you can see the dog's whiskers and you actually see holes in the leaves that he's sitting amongst. And on the right side, you lose that detail and the leaves tend to look a little bit more whole. And this is not meant to be, I guess, the extremist of examples. But the image on the right was also compressed essentially to fit on a cell phone. And so somebody had mentioned that at one point as being an issue. I don't know that it really is, but another key issue that Dr. Badano will talk about later on is that it's possible to capture an image with one pixel arrangement and display with another.

So, for example, a monitor typically has 72 pixels per inch and displays maybe 1680 x 1050 pixels, and this was the arrangement that I had actually used when I had originally prepared this presentation. This gives you a 3.2 to 2 ratio of width to height. The commercial camera example I gave you is the 21.1 megapixel camera, has just under 4,000 pixels per inch, and it displays or collects images at 36 x 24 or 3 to 2 ratio. So somewhere along the line, there's got to be either extra processing or loss of information or distortion of the image in order to make the capture fit with the display.

So really to summarize, whole slide imaging is the capture, storage, and display of digital images of glass slides. There are possibilities for

variability in each of these steps, both the image collection, the storage, and the display, and really what we're here about is the idea that standardization is needed for us to really be able to evaluate these devices as they come forward.

So I thank all of you for your attention.

DR. ADCOCK: Thank you, Dr. Lorick. There will be time later for the Panel to ask questions of the FDA presenters in the afternoon.

So we are going to go ahead and have Dr. Badano speak to us. Dr. Badano is a Research Physicist from the Division of Imaging and Applied Mathematics, Office of Science and Engineering Laboratories. His presentation is on the display of digital images, hardware and software.

DR. BADANO: My job here is to finish up the line up of speakers in the morning with a look at display devices, and I do that by, I have an introduction on display technology, and I'm going to rely on the experience that we have at FDA in terms of the transition into digital radiology. I'll talk about characterization methods that we have available today for display devices, and I will also talk about those methods that we do not have, and I'll talk about some remaining issues, particularly with applications to WSI systems, and I'll summarize.

So let me start by, maybe one of the most important points in my talk is that displays are components of the imaging chain. In the bottom there, you can see there is cartoon representation of a digital radiology

system where you have acquisition, processing, storage, transmission, processing again, and display, and it has been documented that for some applications in digital radiology, the display can be the weakest link in the chain. If you have a display that has poor quality, it can cause incorrect or inconsistent decisions, increases the variability among readers and within readers and leads to longer reading times.

So if you think of the display device as a component of the WSI system, then we can think of having a regulatory path that will go by component, and you will hear more about this tomorrow in Dr. Myers' presentation on the history of digital radiology at FDA. It also allows you to do modular system improvement. So you can plug and play different components of the system as you go along in technology improvements while raising some questions about interoperability of these components.

However, all of this is only possible if standard characterization methods are used to demonstrate or to validate the performance of the different components. In this case, we're going to be talking about displays and as you see, this is not a very good display. So some of these images will not show well, but let me first speak about the technological innovations that we've seen around the year 2000 in digital radiology. We went from the cathode ray tube, a very mature and old technology, to the flat panel technologies, liquid crystal displays of different formats, 3 megapixel, 5 megapixel, 9 megapixel, and we seen also the emergence of some other

display technologies including portables and flexibles.

More recently we've seen that this technology progression continues with very large format displays, 8 million pixels in a single panel, multiprimary displays, which instead of having these RGB dots have 5 dots with different colors which allows you to represent more colors on the screen compared to the RGB panels. We have flexible displays. We have mobile displays, electronic papers. Stereo displays are a very popular right now in gaming applications and will at some point go into the medical imaging arena, and we have high dynamic range displays as well. I have not included high dynamic range topics in this talk because of time constraints, but if there's any question regarding that, we can talk about it in the afternoon session.

This chart is to make the point that the technological innovation in displays requires technology independent standard characterization methods. In other words, if a system comes in with a different technology, we have to be able to reconcile the testing methods so that we can understand what the image quality of those devices is.

Fortunately, in some sense, most of the displays that we deal with today in medical imaging and in most other markets are based on the liquid crystal technology, which is a transmissive technology. You have a backlight of some sort, and the light goes through a stack of several components, and at the end you get an image in the screen. This is of particular importance because the quality of the image depends on what

comes out of the backlight, and what is differentially absorbed by the different layers particularly color, as you will see at the end of my talk, is affected by the fact that this transmissive technology doesn't represent the same image if you go at a different viewing angle with respect to the normal incidence, something that you've all experienced with viewing angle problem in LCDs.

So let me talk a little bit about the characterization methods that we have and we've used in the digital radiology field. There is a very comprehensive cookbook of measure tests that VESA has put out, and you have the references at the end. This is broadly applicable to all display markets and applications.

ACR has published a few years ago a simplified version of requirements for practice in radiology which is useful. A more comprehensive approach has been taken by the American Association of Physicists in Medicine, Task Group 18, which published a very long document of testing procedures for characterizing display performance in radiology.

The same document was the basis of a very recent IEC document that will be published in December. This is actually the first international standard on image displays for medicine for medical imaging in this case. I will talk a little bit more about these two efforts which I think are key in this area.

TG18, as I said, is a long document that provides testing

methods for a number of aspects of display quality. You will see, some of you will recognize this as an evolved version of the SMPTE pattern, or S M P T E pattern, which was used by TV manufacturers in the early days to determine display quality for TV purposes and has a variety of test patterns regarding different characteristics, luminance response, noise -- resolution, uniformity. There's also a few what they call clinical images which are used for checking the quality of the display before a reading session.

In the IEC, we have also a variety of testing methods including gray scale luminance, chromaticity, pixel faults or pixel defects and so forth. You'll see that there are visual evaluation methods that can be done by a trained operator of the systems, and there's quantitative methods that require special instrumentation.

The visual tools again use some form of test pattern, and particularly here, well, because this is not a very good display, you will not be able to see the words quality control in these three panels here, but a good performing display for radiology should be able to very clearly discern these letters on this background, and it has a number of features for different display artifacts.

In terms of the quantitative tools, we've recognized the possibility of using different instrumentation to characterize displays. There's telescopic approaches or what's called as the embedded systems or the puck system that you put on front of the display to measure the display

characteristics.

All of these standards, which I have to say apply only to gray scale displays, have been used by FDA in a recent guidance document to specify a list of parameters that sponsors of digital radiology systems might want to consider when submitting their applications, their submissions. So you have descriptions of the hardware and software, sizes, dimensions, luminance response in a number of ways, reflections, veiling glare. You can read most of them. I want to point out perhaps that some of these like pixel defects are always relating to the AAPM and the IEC standard as standard methodologies for approaching these problems.

Okay. So everything that I've said so far is related to gray scale displays.

What we do not have at the moment is an extension to color displays which is needed for WSI. So here are a number of documents that we can rely upon for some aspects of color, where there's been a lot of work in graphic arts. There's an ISO document but this is a reflection mode standard, meaning that the color is determined from, there's an illumination source reflecting the sample and then measured. That's not the case for displays. Displays are self-luminous devices, and there are some documents that we can use particularly looking at the instrument characteristics needed, comparing different types of colorimeters and several organizations, including ICC, International Consortium for Color, have raised a number of caveats for

these when they are applied to digital displays. So they're not the end word on this topic.

There's also another ISO document that's just been released a few months ago, but that's for color proofing. So again it's comparing color in the display to the color of a printed or a physical sample.

In summary for this part, we do not have a practical standard for measuring the physical characteristics of displays for this type of systems yet. Some of the topics that we might want to consider are color mapping from source to display, display uniformity, resolution and noise, in the similar way that the gray scale standards have defined those, color consistency and calibration, color luminance in terms of what meters, what instruments are needed to make that determination, and also the color response with angle which I'm going to talk about later on.

Okay. So some of the remaining issues. These have been covered in some way or another during the first talks this morning, but I'll tell you what happened in radiology. We had digital detectors that were able to acquire from 2 to 9 million pixels and are often much larger than the display formats that we have in radiology that vary from the 2 megapixel to the 5 megapixel.

The initial approval of some of these systems, particularly for digital mammography, relied on the very high performance displays that had quite similar number of pixels between the detector and the display. And

more recently, displays with lower image quality, in this particular case, lower array size, are being considered based primarily on physical measurements. So that might be a scenario that we can think for these WSI systems.

Another aspect that's well understood not only in the community but outside is the fact of ambient lights. We do have the lights dimmed here in this Panel because of that, and in digital radiology, reading rooms have controlled ambient illuminations that vary between 1 or 2 lux to maybe 10 or 20 depending on where you are across the different clinics. There's only indirect lighting allowed. So if you have a display and there's another workstation on the other side of the room, they cannot be facing each other at a certain distance because that will be direct lighting and would interfere with image quality of the displays. And such environment of controlled ambient light reduces the variability and give us some assurance that the device is performing as it should.

The final remaining issue I'm going to talk about is color. I'll use the term of color gamut to go through this. Color gamut is the range of realizable colors by a device, and if you look at one of the charts, this is just a graphical representation of the colors that the human visual system can see which are inside this shape in particular units. It doesn't matter at the moment. And if you think of a particular object that you're imaging as having colors within that white triangle, and then you think of the display device that can only put out a reduced gamut which is within this black triangle, it means

that some of these points will have to be transferred or mapped into the availability colors of the display.

So one of the questions that we could pose ourselves is what fraction of color that exists in pathology slides through the light microscope would be outside of the camera and display gamut? You can see in this on the left. This would be a compression method. On the right, you would see a clipping or saturation method where all the colors that are outside of the black triangle are actually reduced to single points in the boundary of the triangle. So you are eliminating information in that sense.

Most displays today for commercial uses are designed to have a color intent that is either for TV or for photography, which means that the mapping is preferentially done to represent some colors very well and not so well some other colors. For instance, skin color is very important for photography and TV applications.

There are four ways of doing the mapping, perceptual, saturated is used very much in CGI and graphics, of course, very artificial colors. Relative for logos when you have to combine two colors and you want to see the difference, but perhaps absolute colorimetric mapping should be performed in the WSI systems.

And the final slide goes into the topic of viewing angle. In LCDs, we've shown that, and it's been recognized that, the image that you see at different angles changes as you move it around. In particular, we've actually

measured the changes in the color as a function of viewing angle, and I apologize, this display is very small, but this point here is the color coordinate of the red emission head on perpendicular to the display, and if you go 15, 30, or 45 degrees, you see that the red actually moved to the west if you want in this case.

For the black here, this is the measurement on axis, and when you move to an angle, it goes northwest in this case. So you can see that not all the colors move in the same direction, which means that it would be possible to have color inversions at different viewing angles.

So I'll summarize with a few more thoughts that I have not gone into details right now. We can go in the question session.

I've talked about display as a component but, of course, there is a whole system that needs to be considered, and therefore color management using perhaps ICC profiles, which are standards recognized by the digital photography and color industries, is appropriate.

We will talk again about reference targets in the afternoon, but this could be reflective, transmissive, or actually digital patterns, and there are uses for most of these in characterizing WSI systems.

And then I haven't talked about the calibration patterns, stability of a single device, the QA/QC procedures, and I have not talked about the consistency between different displays in presenting the same information to the readers.

So the main two points of my talk, if I can remind you of those, is that display systems are components of the WSI systems that need to be considered, perhaps as individual components and as a whole system testing approach as well. And then there are some areas in characterization of displays that will be very useful to evaluate the performance of the devices and therefore facilitate the FDA review.

That's all I have. Thanks very much.

DR. ADCOCK: Thank you, Dr. Badano. We will now adjourn for lunch, and we will reconvene in this room at 1:15. Please take any personal belongs with you at this time. The meeting room will be secured by FDA staff during the lunch break. You will not be allowed back into the room until we reconvene. Thank you.

(Whereupon a luncheon recess was taken.)

AFTERNOON SESSION

(1:20 p.m.)

DR. ADCOCK: It's now 1:20, and I would like to call this meeting to order.

MS. MAGRUDER: I'd like to just state for the record that Dr. Descour is a major shareholder in DMetrix and on the Board of Directors for the company. And Dr. Balis is on the Technical Advisory Board and has shares in Aperio. Those were left out this morning. Thank you.

DR. ADCOCK: Thank you, Louise.

We will now proceed with the open public hearing. Public attendees are given an opportunity to address the Panel to present data, information, or views relevant to the meeting agenda.

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 of this meeting. For example, this financial information may include the company's or a group's payment of your travel, lodging, or other expenses in connection

with your attendance at this meeting. Likewise, FDA encourages you at the beginning of your statement to advise the Committee if you do not have any such financial relationships. If you choose not to address this issue of financial relationships at the beginning of your statement, it will not preclude you from speaking.

The Panel will be given an opportunity to ask questions of the public presenters at the conclusion of the open public hearing. If recognized by a Panel member, please approach the podium to answer questions.

I would like to remind public observers at this meeting that public attendees may not participate except at the specific request of the Chair.

Prior to the meeting, we received seven formal requests to speak at today's open public hearing session and can only accommodate these individuals. As I call your name, please come forward to the microphone. We ask that you speak clearly into the microphone to allow the transcriptionist to provide an accurate record of this meeting.

You have six minutes for your remarks. When you begin to speak, the green light will appear. The light in the timer will turn yellow to warn the speaker when there is one minute remaining. In the interest of fairness to the other participants, we ask you to conclude your statements within the six-minute timeframe.

The first speaker will be George Netto, reading a statement

from Dr. Juan Rosai. Dr. Netto, please come forward to the microphone. Again, we ask that you speak clearly to allow the transcriptionist to provide an accurate transcription of the proceedings.

DR. NETTO: Good afternoon. My name is George Netto. I'm an Associate Professor of Pathology, Urology and Oncology at Johns Hopkins. I'm reading a statement on behalf of Dr. Juan Rosai. The statement reads as follows:

Dear Sirs:

My name is Juan Rosai. I am an M.D. and a senior surgical pathologist with a long experience in American academic (and lately private) medicine. This includes pathology training under Dr. Lauren Ackerman at Washington University in St. Louis, Director of Anatomic Pathology at the University of Minnesota in Minneapolis (10 years), Director of Anatomic Pathology at Yale University in New Haven (6 years), Chairman of the Pathology Department at Memorial Sloan-Kettering Cancer Center in New York (10 years), Chairman of the Pathology Department at the National Cancer Institute in Milan, Italy (5 years), and currently holding the dual position of Senior Diagnostic Pathologist at Genzyme-Genetics, New York, and the directorship of the International Center for Pathology Consultations at the Centro Diagnostico Italiano (Italian Diagnostic Center) in Milan, Italy.

I am the author of the textbook, *Rosai and Ackerman's Surgical Pathology*, now in its ninth edition and widely regarded as the premier

publication in the field. I have also been the Editor-in-Chief of the Third Series *A.F.I.P. Atlas of Tumor Pathology* and senior author of two of the fascicles of that series, *Tumors of Thymus*, Second series; *Tumors of Thyroid Gland*, third series.

I am currently collaborating with Aperio in a series of didactic projects, one of them co-sponsored by the United States and Canadian Academy of Pathology, USCAP, the largest pathology organization in the country. I'm not a shareholder or consultant for Aperio, and I have no other official ties with them.

For the past 25 years, my diagnostic work has been limited almost exclusively to the examination of consult pathology material submitted to me by pathologists, clinicians, and increasingly the patients themselves. In nearly all of these cases, I receive a set of glass slides and a brief summary of the clinical history, supplemented, if indicated, by paraffin blocks, x-rays, CT scans, MRIs and other pertinent material.

I have explored many years ago, in collaboration with Dr. Stephen Erde, at Cornell University, and Professor Vincenzo Eusebi, from the University of Bologna, Italy, the possibility of performing part and eventually most of my consultation work, especially the one originating from overseas, using what at the time was a rather primitive technology and reported on our early encouraging efforts in the *Journal of Human Pathology*, Volume 28, page 13 to 16 (1997).

I have followed with increasing enthusiasm the impressive advances that the technology has undergone in recent times, until reaching a level such that I believe it matches and, to some extent, surpasses the capabilities of the traditional examination of glass slides under the microscope.

With the best digital instruments currently available, the image resolution (the absolute key feature in a microscopic evaluation) is just as good, if not better, and the capability of manipulating the image (moving to different fields, changing magnifications, changing focus in some of the models, etc.) is certainly easier.

There are actually some aspects of the procedure that are better carried out with digital images than with traditional slides, such as examination of the material at very low (panoramic) magnification, the simultaneous examination of low and high power appearances of the same field, the side-by-side comparison in the same screen on the images of the problem case with known standards, and the capability of quantifying the findings, a procedure that according to some authors will finally elevate microscopy to an objective and reproducible technique. This includes the precise measurement of cells, nuclei, nucleoli, depth of invasion, etc., and the quantification of positive cells in special preparations such as immunohistochemistry, as opposed to the time-honored but highly imprecise eyeballing done currently by most pathologists, their occasional denials

notwithstanding. I don't agree with that. I'm kidding.

Along these lines, I concluded recently a pilot study with three prominent pathologists practicing in Lima (Peru), Buenos Aires (Argentina) and Sao Paulo (Brazil), respectively, in the course of which these pathologists scanned the slides they had selected for my consultative evaluation, sent them to me, received my diagnostic opinions and concluded that my diagnostic accuracy was essentially the same as when, in a second step, I examined under the microscope the glass slides of the corresponding cases.

Being that I receive consultations from many parts of the world, I particularly appreciate what I regard as a big plus of the technique, i.e., the speed of the consultation process (measured in minutes rather than in days or sometimes weeks) and the elimination of hazards such as loss or breakage of the material in transit, or significant delays at the Customs Office of some countries (which we are experiencing with increasing frequency) on the grounds that the slide represents "biologic material" and is therefore incorrectly classified as "potentially hazardous" material, with all the logistic problems, delays and added costs that this decision implies.

Further advantages of the digital technique are the fact that there are no limitation to the number of consultants that one may wish to engage, and the capability of viewing simultaneously the same field and to discuss the findings from several independent stations scattered in a large geographic area.

Whereas as my work, as above stated, is almost entirely devoted to consultation material, it should be obvious that the technique is just as well suited for primary diagnosis, the assumption being that the "routine cases" are, on the whole and by their very nature, easier to interpret than cases selected for consultation because of their complexity. Also not to be underestimated is the fact that the images on which the diagnoses are based can be stored indefinitely in secure servers without deterioration and are not subject to the perils of glass slides, such as misfiling, drying out and breakage. Needless to say, there are legal and logistic issues that need to be worked out by the properly qualified individuals before approving this procedure. I would simply conclude by saying that from a technical and scientific standpoint, I am thoroughly convinced that a diagnosis made on the basis of a well-prepared digital image of a representative whole section is just as informative and accurate as that performed by using the time-honored examination of a glass slide under the binocular microscope.

I hope you will find these comments of mine useful for your discussion and decision. It goes without saying that I would be gladly available for any further discussion of this matter through any venue you might like to choose.

I would like to conclude this letter by expressing my appreciation to the FDA for considering my opinions and views on such an important matter, which I have no doubts will revolutionize the field of

pathology, if it not doing that already.

Sincerely yours,

Professor Juan Rosai,

Centro Consulenze Antomia Patologica Oncologica, Centro Diagnostico Italiano.

Thank you.

DR. ADCOCK: Thank you, Dr. Netto. Our next speaker, Mr. Fatemi, please, or Ms. Pardon me.

MS. FATEMI: Good afternoon. Thank you to the Panel for hosing the meeting today. We appreciate the fact that FDA recognizes there are issues associated with the diagnostic applications of digital imaging that are different from other applications and perhaps unique to light microscopy for pathology.

I am Mahtab Fatemi, of counsel with Longwell and Associates, in Palo Alto, California. We have had a number of clients who are developing and commercializing devices for digitizing microscope slides.

One of the issues we have seen in the past is that there has been an emphasis on specific counting applications rather than a broad indication for replicating, reading and storing image data from microscope slides for the purpose of diagnostic pathology.

As you probably know, there is a standard being developed by DICOM, Working Group 26, titled DICOM supplement 145, Whole Slide

microscopic image IOD and SOP classes. We would like FDA to seriously consider how this standard could be developed into a Class II control once finalized. The standard is currently in draft. Could FDA join this working group or appoint an official delegate to this group?

Adoption and consistent application of a standard would allow for a reproducible level of accuracy, with the fidelity of the image being the main objective. Currently the 145 standard is addressing standardized amount of information (what is to be stored, not just patient ID, but what kind of tissue, what kind of stain and staining procedure and what type of microscope); the dimensions of imaging (the Z planes), size of field, and the need for changing magnification; and lastly, multi-spectral imaging (more than just H&E).

Among other important areas of standardization, first the issue with how to gauge the accuracy of the image produced should be discussed, and we look to this group for help. Additionally, are there quality parameters that need to be included in this standard, or are these outside the scope of the document?

This kind of application cannot be addressed by pathology consensus working group testing because the type of information is too broad to evaluate in feasible clinical studies. But rather, we need a standard that ensures that the imaging parameters are consistent and lead to image fidelity. Diagnostic accuracy can be assured by adherence to a standard in conjunction

with the appropriate calibration and quality control measures.

We would like to ask FDA and CLSI to join this working group and work towards developing DICOM 145 as a Class II control. This would provide clear definition for manufacturers to develop products and ensure consistency and accuracy of imaging data that will ultimately result in more efficient patient care. Thank you.

DR. ADCOCK: Thank you. At this time, would Mr. Soenksen please approach the podium.

MR. SOENKSEN: Thank you very much for the opportunity to address the Panel. My name is Dirk Soenksen. I'm the President of the Digital Pathology Association. I'm also a shareholder in Aperio.

I wanted to first outline to you the mission of the Digital Pathology Association. So there's some slides to be viewed. It's focused on really educating people interested in digital pathology, creating awareness in digital pathology with a goal to accelerate adoption. So we are focused on educational initiatives, defining best practices, influencing and setting standards, and also organizing an annual conference called Pathology Visions.

There are 17 companies, many of whom are in this room, who are members of the Digital Pathology Association. We also have leading pathologists from clinical, biopharma and research institutions, and will soon have members that comprise government and academic institutions as well. So it's the first industry association recently formed focused in digital

pathology.

Digital pathology, as you learned in the technical sessions earlier, comprises really -- slides, managed slides, visualize them and do image analysis. Digital pathology, in conjunction with digital radiology and digital cardiology, will provide in the future for a more complete electronic medical record and for an enterprise-wide imaging solution for healthcare institutions. This is what the focus has been today, and this is I think how the stage was set from the FDA.

Digital pathology will also become critical in personalized medicine as a platform for tissue based quantification of biomarkers. So the relevance to biopharma is quite important with respect to how the FDA views the clearance of digital pathology.

At the same time, digital pathology will in the future be used to automate many tasks that pathologists do today manually, finding rare events, finding microorganisms, and so forth.

So the reason I'm showing this slide is to suggest to this Panel that the scope of digital pathology is quite broad and actually is broader than just simply making a diagnosis on a monitor.

Specifically, we've seen 510(k) clearances as we've talked about for digital IHC, and we've seen clearance for the manual read of a digital slide for PR and HER2 and, of course, the diagnosis on the monitor is what we're talking about here.

I wanted to just broadly outline what's happening in various areas. In education, medical school students are trained using digital slides today. Many medical school students select their residency programs based on whether or not a particular institution offers digital pathology. Vast amounts of anatomic pathology courses that provide CME credits use digital slides, and as we've heard today, the exams given by the American Board of Pathology uses digital slides today. So the reality is that the next generation of pathologists are being trained today using digital pathology.

In biopharma, digital pathology is used in all aspects during discovery by all big pharmas and all the major contract research organizations to streamline their drug discovery process. Applications include imaging analysis, peer review, and the global management of histopathologic data. The National Institute of Environmental Health Sciences and the National Center for Toxilogic Research used digital pathology today. Digital pathology today is in the critical path of drug discovery. Diagnosis on a monitor will enable the use of digital pathology for GLP studies, which is very important.

On the clinical side, many academic medical centers and some hospitals use telepathology for selected applications, frozen sections, archival and retrieval, formal and informal consultations, quality assurance, risk management for tumor boards, and other applications. Major reference labs are currently using cleared applications of digital pathology to provide improved turnaround times. Digital pathology is in use for routine clinical

practice outside the U.S. There are institutions outside the U.S., particularly in countries like Sweden, that are 100 percent adopting this technology for primary diagnosis. They've done this for many years. They've run their own validation studies, and they have found that concordance between glass slides and digital slides is perfectly adequate.

Digital pathology is without a doubt the future of anatomic pathology, and the ability to promote the use of this technology for diagnosis on a monitor will enable improved quality and better patient care.

We hope and urge this Panel before we conclude this meeting to actually drill down on the appropriate regulatory pathway for this technology. Specifically, whether this is a 510(k) or there's going to be a PMA. The implications for the industry, if a PMA pathway is selected, are actually significant because it will slow adoption, and it will make the hurdles so high that it will be very, very difficult for the majority of those vendors and industry to get over those hurdles.

So I think a healthy debate about what the appropriate regulatory pathway should be and possibly recommendations to FDA to provide a view on that is really essential and something that we hope will be undertaken by this Panel before the end of tomorrow.

Finally, I wanted to just make you aware of a conference called Pathology Visions, which is a conference dedicated solely to digital pathology. It is hosted by the Digital Pathology Association. They will be meeting in 2010

in San Diego. We expect 500 attendees, multiple tracks, and we would be delighted if there was representatives from the FDA who could come speak to all of the stakeholders in digital pathology at this conference. Thank you very much.

DR. ADCOCK: Thank you. At this time I'd like to invite Mr. Eichhorn.

MR. EICHHORN: I'd like to thank the Panel for allowing me to make a few remarks. My name is Ole Eichhorn. I am the Chief Technology Officer of Aperio. I'm employed by Aperio and a shareholder.

So I'd like to make a few remarks about measuring accuracy in digital pathology. I know this is something the Panel is planning to consider in depth tomorrow.

There are, in our view, several aspects to measuring accuracy. First, it's very important to develop a method to record diagnoses. Then it is also important to develop a method to compare two recorded diagnoses, and the final and key component is to develop a method to determine the ground truth diagnosis for a cohort of test slides.

With those three capabilities, it's possible to record diagnoses, record truths, and then compare diagnoses to the truth in order to measure accuracy.

So now I'd like to make just a few more remarks about these key aspects. As you know, pathology reports are complex. They're English

text with medical terms, and different reports may correspond to the same diagnosis. This makes the complexity of recording a diagnosis difficult compared to a prognostic indication such as immunohistochemistry scoring. There are nuances in reports that vary by lab, by region, by training, by the age of the pathologist, and so on. So it's a difficult problem.

We would like to suggest that it's important to record diagnoses in a structured way, that's tissue and preparation specific, and a good starting point for recording diagnoses is to use the College of American Pathologists checklist. These checklists have been developed with many, many man and woman hours of work, very carefully, and they provide a really good starting point for recording a diagnosis.

From there, it's important to record all the significant features that would be noted on a report that affect patient care, and the diagnosis can be cross-checked to ensure that it's consistent, and this can be used for recording any diagnosis that's tissue and preparation specific irrespective of the method, whether there's a device being used or conventional microscopy.

Once two diagnoses have been recorded in this structured way, they essentially constitute a tree of choices, and the two trees can be assessed node by node to assess the severity of each difference, and the differences can be assigned a different comparison result based on the impact on the patient. On any given comparison, there can be no difference or it can be an incidental difference. It can be a significant difference, which would be

something that two pathologists would ordinarily agree upon, and then most importantly, any difference that affects patient care is a clinically significant difference. And as two diagnoses are compared, the most severe difference on any part of the diagnosis becomes the severity of difference between those two.

So with this method in hand, it's possible to take two diagnoses that have been recorded in a structured way, perform a very rigorous comparison, and then come out with a result that's useful for statistical analysis.

Determining the ground truth is quite difficult. As all of you know, this is a key aspect of conducting these studies, and we hope that there is good discussion on this topic in the Panel so that you can guide the FDA so that the FDA can guide vendors. Pathology is inherently subjective. A lot of cases are obvious, where most pathologists would agree, but difficult cases are typically diagnosed via consultation. And as was noted in some of the discussion this morning, pathologists generally know when they need a consultation from a colleague or re-cut of a slide, or they sort of know when they know.

We suggest that a good way to record the ground truth in a structured way is to convene a panel of expert pathologists, diagnose a cohort of slides using conventional microscopy and then resolve any differences that there may be using consultation in the same way that pathology is conducted

in the normal clinical setting. And then the consensus is recorded using the method to record diagnoses as I had just discussed.

Combining these yields a process to measure accuracy. A diagnosis from a cohort can be recorded, the ground truth can be determined and recorded, and then the diagnoses can be compared to the ground truth, to the yield, a very rigorous way of measuring the accuracy.

So thank you very much. I appreciate your attention.

DR. ADCOCK: Thank you. Mr. Dunstan.

MR. DUNSTAN: My name is Bob Dunstan. I've been an anatomic pathologist for 30 years, and I've asked to speak in order to emphasis the transformational impact whole slide digital imaging will have on my profession.

Anatomic pathology has long played a critical role in diagnostic medicine and research, and I think with technology, we can do a much better job.

I don't know if you know it or not, but anatomic pathology, we're sort of like the Amish of the biomedical community. I say that because how we process tissues and how we examine tissues really hasn't changed in a century. Okay. If you look at our microscopes, with the additional of electricity and binocular lens system, they're the same. Okay. In fact, a veterinary pathologist can read a FDA -- going to FDA GOP study using this microscope. Okay. In contrast, look where radiology has gone in a similar

period of time. Okay. And the point to realize is that after x-rays, which are now re-digitized, everything else has required the acceptance of digitization. And I believe that if we accept digitization, the similar type of changes will occur in our profession.

The point I would like to emphasize is basically it will have a profound impact. First of all, it will improve efficiency and accuracy without a doubt. I think accuracy is already stated. When you can hold the previous and the current biopsy right next to each other, side by side, that will improve accuracy. I think it will open the silo, which is largely considered to be a good portion of my people outside of pathology, but the thing I'd like to make the greatest emphasis on is that it will turn a report from being descriptive to descriptive and quantitative.

Pathologists are skilled at recognizing and describing microscopic patterns that correlate with disease process. When a discipline is descriptive, there is a great tolerance for how a sample is handled. This is what has happened in anatomic pathology where there is no standardization for how a tissue is fixed, how it is processed, or how it is stained. I would imagine if you ask each pathologist in this room how they get to a slide, every case will be different, and this is not a problem as long as the data we present remains descriptive. What is currently happening with pathology is that with the emphasis of immunohistochemistry and the need for biomarkers, more and more pathologists are asked to correlate the pattern with the protein,

and this is where problems are becoming.

So, for example, see the problem is, okay, is that eyes are much better for identifying pattern than for quantitation. For example, let's look at the Hercep Test. Okay. This is a FDA approved test, okay, to quantify the expression of HER2 in tissues. On the surface, nothing could be easier, okay. You stain a slide, and if greater than 10 percent of the tumor cells, the breast cancer cells have the membrane stained at a 3+ level, the patient can get Herceptin as a therapy. However, even with the most experienced pathologists and laboratories, the error rate is greater than 20 percent and exceed 30 to 40 percent in less regulated environments. Okay.

And the point to be made is, it's never going to get easier than this, okay. Herceptin, the HER2, is amplified due to gene amplification, and this is a difference on a HER2 density on a 1+ score versus a 3+ score. It's a 70-fold difference, and I know of no marker, no surface antibody, being developed in tumors today that have this wide a range between 1+ and 3+. So we really have to get better.

The point is, as long as we remain totally on basically 0, 1, 2, 3+, we are always going to be bottom feeders in terms of biomarker development. Okay. There was a recent article by Cummings, et al., in the *British Journal of Pharmacology*, and what we're doing is considered in terms of rigorousness of biomarkers as only qualitative. It's not even considered quantitative. Okay. The best we're currently doing in pathology today, okay,

is having a computer. So what we have is a continuous response variable, and that's only considered quasi-quantitative.

So for us to move further on, okay, we're going to have to develop reference standards, and this is where digital whole slide imaging comes in. It's the very, very first step, okay. So basically without digitization, we can't quantitate, and without quantitation, we can't standardize. Okay. And that's why we talk, there are no gold standards in pathology because we can't digitize or quantitate, and I think that if we really want to sort of play a role in the generation, and that's the future, okay, of basically prognostic or predictive or pharmacodynamic biomarkers, okay, we're going to have to standardize, and we're going to have to develop reference standards. And I propose that if we can't, then really the value of correlating pattern with protein is lost, and I think that's a tragedy for pathology and a greater tragedy for our patients.

Thank you.

DR. ADCOCK: Thank you. Mr. Cartwright.

MR. CARTWRIGHT: Okay. You didn't get my presentation. My name is Gene Cartwright. I'm CEO of Omnyx. So I'll just address my words to you by leafing through my slides, and I'll make a few key points. You have my slides. I don't know why my presentation isn't here. At any rate, let's move forward.

So let me first tell you who Omnyx is just very quickly. We are a

joint venture between UPMC and GE. We are focused, and that's the point of this first slide on clinical pathology, on digitizing clinical pathology. There are a number of challenges as you can read here, obtaining excellent image quality at speed because as you know, the pathology laboratory, the clinical pathology laboratory is somewhat like a factor environment, and achieving think speed image streaming.

I'd like to add a point here, and that is that we focus on, we split up a 20X image into maybe 400 tiles, and we focus on each one of those tiles, taking multiple planes of focus and then in calculating an ideal focus plane for each tile. My point in saying this is not so much to say that we have the best technology. It's just to show you that technology continues to move forward.

If we go to the next page or the next slide, I just highlight a number of the improvements that we would gain with digital pathology that have been spoken about quite a lot before. So essentially we'll just go ahead and move on, but I do want to emphasize, as Mike Becich did, that there are a number of significant quality improvements that this technology can provide.

I have some observations and questions on the next slide about how to design a clinical study, and I agree in many aspects with Ole Eichhorn's comments. But, first, obviously we have to compare the pathologist's reading of a digital image to the pathologist's reading of a glass slide. I would also agree with Mike Becich who responded earlier that there are certain areas

that we would exclude, cytology, aspects of hematopathology, samples that are greater than 12 microns in thickness, and as a note, IHC is addressed already as a separate approval path. We also are not advocating the elimination of glass slides. We agree with Mike Becich's comments on that.

If we go to the next slide, I have just a few more questions. You can assess the performance of this technology in a number of different ways, optically, clinically, and potentially with phantoms. We would not recommend a simple optical assessment of performance. A simple micron per pixel specification would not really be sufficient in our mind. So clinical is probably what we'd recommend. There are a number of questions you can read here that we would like to consider. What are the criteria for sample selection? Is it a very broad range of samples? Is it difficult samples? How do you define difficult samples? What is the statistical rationale? I know that will be a key question. What's the comparative performance standard that must be achieved? So, you know, is it 90 percent, a diagnostic concordance? Is it 100 percent? What's the level of uncertainty?

And then I think importantly is the performance judged with digital alone or is a pathologist allowed to ask to reflex to a glass slide as they will do in the real world? I think it's an important consideration.

From an artificial standpoint, using phantoms, I think the issue there is we're all in favor of looking at phantoms. That's not an issue from our standpoint. However, one of the most important aspects of doing digital

pathology is the ability to focus. So if we do use phantoms, we'd recommend that they have some pretty significant topography to really test the whole slide imager.

The next slide just talks about essentially what is the gold standard? Many people have asked the same question today, and I think Ole's comments I would agree with in many ways, that there has to be a consensus, looking at the glass slide to really get the initial gold standard.

We would also highlight two different ways of looking at the clinical performance. One concordance that we've already talked about but the second Mike Becich mentioned a little bit which is to identify features which are important diagnostically but at the edge of detection of this technology, and if we can then reproducibly detect those features and show that we reproducibly detect those features, that would be another very good indication that this technology is ready.

If we go to the last slide, there are only two or three points I'd like to make. One is on image compression. It's very clear that you can over-compress a digital image and get an image that's not useable. Our recommendation is, though, that let the companies decide what the compression should be but then focus us to do the clinical studies to prove the performance of the product. Don't set a requirement on compression by itself. That would be our recommendation.

Again, we are recommending to allow the pathologist to reflex

the glass as I mentioned earlier.

And there are two risks. One I mentioned here, which is how can we best assess the risk that a pathologist will not know when the digital image is inadequate and when they need to reflex to the glass slide. That's been brought up earlier.

I would also agree with Dr. Gilbertson's comment about proving that we can scan all the tissue that's on a slide.

Those are the two risks that we see. Otherwise, we think that by looking at the clinical performance, we can prove that we prove adequately.

So thanks for your time and thanks for arranging this meeting.

DR. ADCOCK: Thank you. And our last speaker, Dr. Fonyad.

DR. FONYAD: Okay. Thank you very much. My name is Laszlo Fonyad. I'm a pathologist, Semmelweis University, Budapest, and I have to disclose my relationship is really -- where I'm an advisor of the educational software development.

Today I'd like to talk about one of my studies measuring diagnostic accuracy using digital slides and the routine. In the first couple of minutes, some background and some -- about why we designed our study and the materials and methods we used and our results, and because of shortness of time, just some selected results I mention and, based on our results, some recommendations.

Okay. So even though that the spreading of digital microscopy is worldwide, it's unquestionable, we think that the real revolution of digital slides are still waited for. So we wanted to estimate the major causes of dislike and/or dissatisfaction that could explain the mistrust in digital slides.

Our questions were can we define a list of samples according to the origin and ties where digital microscopy is sufficient to use? Is it possible to estimate types of error resulting misdiagnosing cases? And does the pathologist's interpretative skills affect the results of using digital microscopes?

Altogether, almost 300 cases were enrolled to the study. It was semi-randomized, very simple cases like appendicitis, colitis was sorted out, and almost 100 and 500 slides were reevaluated by the pathologists.

In the study, seven pathologists were involved. Three pathologists received cases specific to their cases, and four pathologists received non-field specific cases. Then a clinical research form was filled out concerning about scan quality and the diagnostic confidence. And later the incoherent cases were graded, and four types of diagnostic errors were defined concerning whether the incoherency was relevant or not and whether any kind of uncertainty was recorded or not.

Here are some technical results, and the most important here, I guess, is that when a slide was rated poor, in most of the cases, it was because the important areas of the slides were off the focus. Other reasons

of disliking a slide were of minor importance.

Okay. Here are some results, and here we see that it's 2.5 percent of the cases that fall into that fourth type of discrepancy where the incoherency is relevant and no uncertainty was recorded.

One of our questions was whether we could define a list of samples according to the origin, whether original microscope is sufficient to use, and we think that anytime when the incoherency ratio is below the ratio of the reassessed cases are samples where we can use digital microscopes properly.

Here I show you when excluding the non-field specific cases, we see significant decrease of incoherency ratio.

So based on our and other results, we think that the level of diagnostic confidence using digital slides are acceptable, and reasons responsible for diagnostic errors are mostly personal and reflects the competency of the examiner.

Technical reasons, potentially responsible for errors, such as poor color fidelity, blurred image, are detectable by the examiner, and correction of it could be initiated.

However, we think that strict regulations required for the scanning process inserted to the pre-diagnostic phase. First of all, safety of sample recognition, how do we prevent data loss because of incomplete scan, proper glass slide handling, provide accurate slide ID recognition, and

minimalize the chance of breaking glass slides during the scanning processes.

Thank you very much for your attention.

DR. ADCOCK: Thank you. I would like to thank the open public speakers for their presentations.

Does anyone on the Panel have any questions for the speakers at this time?

Dr. Gilbertson.

DR. GILBERTSON: Yes. The question is to Dr. Cartwright.

DR. ADCOCK: If you could approach the podium.

DR. GILBERTSON: In your third or fourth slide, you talk about certain potential exclusions in pathology for whole slide imaging, you say things like slides, samples that are greater than 12 microns, samples requiring more than 40X magnification, cytology. Why?

DR. CARTWRIGHT: Well, that's a good question. It's not because the technology can't do it. The technology can do it. We as a company have decided that there's enough significant benefit on the large majority of pathology that is within or is separate from those exclusions that that's what we're going to approach initially, but technologically, there's no reason that the others can't be done as well.

DR. GILBERTSON: Okay. So you're talking about your company's path?

DR. CARTWRIGHT: That's right.

DR. GILBERTSON: Okay.

DR. CARTWRIGHT: Our company's path. So that would be essentially part of our intended use in a way.

DR. GILBERTSON: Okay. Thank you.

DR. ADCOCK: Any additional questions at this time?

Before we move onto our next speaker, at this time I'd like to allow Dr. Balis to present four slides which were inadvertently left out of his presentation, and these are some images that I believe he'd like to share with us. I'm sorry for that omission earlier.

DR. BALIS: Very briefly, the sequence of snapshots from whole slide imaging are meant simply to be illustrative of some of the problems with the current technology and why it may require then some further expiration of characterization of not just the acquisition technology, that is the CCD, the analog to digital conversion process, and then the storage and also the rendering. And, for example, the very fact that I have a reasonably good contrast and brightness on this laptop screen which differs from that of the monitor is intrinsically illustrative of the challenge with the rendering of digital content, and it's not just histology or cytopathology but really any type of digital content. There needs to be characterization and calibration of such systems.

So even on the rendered image on the screen, which is meant to be illustrative of a problem, which is a lymph node architecture, I think we

can all agree that the image is dark and, moreover, there are additional particular cells, lymphocytes which are so dark, both on the laptop screen and on the rendered screen for projection, to render them not available for diagnostic query, whereas with a microscope, there's always the availability of the temperature control, the filament temperature, if it's an LED, then increase the intensity such that you can increase the dynamic range and effectively be able to visualize high contrast or dark regions and appreciate the texture. Absent texture, you're missing information, and that's simply the point, and given a finite capture range of 256 shares of red, green, and blue, there are going to be instances, unless the system's highly optimized, in which various important subject matter will simply be too dark, and that would be an example of reasonably experienced or appropriately informed diagnostician, pathologist, cytotechnologist would defer to slides from the original media.

Similar lymph node diagnosis again with improperly set color temperature, everything looks red here, whereas in point of fact with appropriate color temperature, this would be an equal balance of hematoxylin and eosin. Again, without normalization of systems, not only can this be due to acquisition but because of improper setting of the color temperature of the projection media. So again, one of the areas in which there is the need for standardization.

And in the specific case of oil immersion microscopy, the ability

to look at subcellular structures, granules, what color, what size, what spacing, in terms of hematopathology. I can see them on the laptop screen. You can't see them on the projected image, and it's not a resolution image in this case. It's a rendering problem, but again since the overall diagnostic process now has digital media coupled to the optical component, any failure, which I highlighted earlier this morning, in the overall chain that ultimately provides an image, any one of those components, including the display monitor, being suboptimal, and you failed the overall mission of providing a diagnostic quality image to the interpreter. And again, shifting color temperature and losing dynamic range, becoming too dark to be diagnostically useful.

So again, simple, simple concepts but of paramount importance in ensuring that what starts out as the superb quality of digital data, the digital acquisition, stays superb every step of the way up to the point of visualization in whatever display technology that is selected. That's really the examples I wanted to show. Simple, simple examples.

Thank you.

DR. ADCOCK: Thank you.

Our next speaker from FDA is Mr. Agrawal, an Electrical Engineer from the Office of Science and Engineering Laboratories, Division of Imaging and Applied Mathematics. His presentation is on preclinical bench testing of whole slide imaging.

Following Mr. Agrawal, Dr. Robinowitz, a Senior Medical Officer and Pathologist from the Office of In Vitro Diagnostic Device Evaluation and Safety, Division of Immunology and Hematology Devices will continue the presentation.

DR. AGRAWAL: Thank you. And good afternoon to the Panel and all the attendees here.

I'm a specialist in optical diagnostic and optical imaging medical devices, and Dr. Robinowitz and I will be presenting some of our current thinking on preclinical bench testing of WSI systems.

The motivation behind this kind of testing is threefold. We wish to obtain quantitative, detailed performance data under well-controlled conditions which will allow both manufacturers and the FDA to perform consistent evaluations and comparisons of device performance. By establishing proper preclinical testing procedures, we have the opportunity to reduce the burden of acquiring and analyzing clinical data. In other words, it may be possible to simplify or reduce the size of clinical studies used to validate device performance. And finally, as more of these systems come onto the marketplace and they remain in use for long periods of time, we need to have some methods to ensure that the performance is maintained. So bench testing provides a straightforward means of postmarket quality control.

There are two main categories of preclinical bench testing. The

first is physical testing of imaging performance characteristics, and Dr. Descour this morning provided some aspects of this type of testing, and I'll be discussing this further, and here we're interested in quantitative measurements of key hardware and software properties, and we can also in this type of testing isolate effects of individual components of the WSI system.

The second category of testing which Dr. Robinowitz will be discussing is pathologists or human observer based assessment of image quality using artificial phantoms or biological specimens, and here we're trying to answer the question, are all critical features visualized by both optical and digital microscopy systems? And this type of testing serves as a bridge to clinical studies.

So now I'd like to get into the first category of preclinical bench testing, the physical testing of imaging performance characteristics.

This is a simplified, generalized block diagram of the WSI system which begins with a slide being passed under basically a conventional optical microscope consisting of the light source, that is filament or arc lamp or LED source, as well as a condenser and any aperture diaphragm followed by the imaging optics, that is a high NA objective, the tube lens or the relay optics. Then from there we have a mechanical scanner, and this mechanical scanner could be on the front end actually scanning the slide or where it is here in this diagram, actually scanning the instrumentation, but the point

here is to examine all of the fields of view on the slide as well as the various focal planes in depth on the slide. Then we have some sort of digital image sensor, a CCD or CMOS camera, to capture the image.

The software then transmits the image from the sensor to the computer as well as performs enhancement and compression of the image and then outputs the image data files.

And finally we have the display which is what the pathologist then views to perform the diagnosis.

So we can look at each of these components individually because each of them possesses properties which can impact the overall performance of the system, and I would like to now go over some of these characteristics from each of these components.

Starting with the light source, we have to be thinking about the temporal stability of the light output, the spatial uniformity and the Kohler illumination characteristics thereof, as well as the color or spectral output of that light source.

With the imaging optics, we have a number of things such as the magnification or NA of the objective, the resolution and contrast inherent to the imaging optics, aberrations such as distortion introduced as well as the depth of field, which is, of course, very critical when looking at 3-dimensional specimens.

The mechanical scanner as I mentioned, we are using the

scanner to scan both laterally as well as in depth. So it's 3D positioning accuracy is essential to understand in terms of the overall performance of the system.

With the image sensor, there are a number of things, such as the electronic noise, the read noise, the dark current noise, for example, the linearity of the sensor's response to light, the dynamic range which we've been hearing a lot about this morning and this afternoon. The spatial resolution, of course, the pixel size involved in the sensor, and the spatial uniformity of the response of that sensor.

With the software, we have things like enhancement and compression that can affect the overall imaging performance, and then finally Dr. Badano went over quite a bit on the display issues. So I just listed some of those here, the color gamut, color accuracy, luminance, the resolution and contrast of the display as well as its spatial uniformity. So again here this is just a sampling of the issues or the properties of individual components which can be important to understand during this type of physical testing.

So now when we look at the system as a whole, there are also characteristics which we can then summarize which really are influenced by each of the individual components or certain individual components, and we can look at in total with the system put together. Again, the resolution and contrast is influenced by the imaging optics, the sensor and the display, the color gamut and accuracy, dynamic range, spatial uniformity. All of these we

saw within individual components, and now it's important to understand how the system as a whole is impacting each of these characteristics.

So as far as conducting this kind of testing, we can imagine some simple paradigms to do this. By replacing the slide with some sort of test target and then passing it through the system, and then one way to do this is simply to examine those files coming out after image processing with some sort of analysis software and compute, a figure of merit, which is a quantitative measure of performance. And in that way we have either characterized all these components together, not including the display at this moment, or depending on the type of test target and the procedure we use, we can actually isolate the effect of an individual component.

Another regime of testing which is a little bit more complex, but it could also be very useful here, is to really look at the system from end to end and place some sort of a light measurement instrument here at the display, that is like a spectro-radiometer or even a camera to look at the display and then from there compute a figure of merit. Now, we really have looked at the entire system in total.

It could be important to obtain statistics during physical system testing, and there are two types of statistics which are very readily obtained during the physical system testing. One is to look at within system variability just by inserting some test target some number of times into the same unit and computing that figure of merit each time and therefore then computing

statistics on that figure of merit, and similarly between system variability where we have multiple copies of some WSI unit and we insert the test target into each of those units and again computer statistics on a figure of merit. So these statistics can be very useful and very effectively computed at this phase.

We've been talking more about the limitations I think on what phantoms can do for us. Most of the ones that are available are currently basically 2-dimensional in nature, but that doesn't mean we still can't obtain useful information at this point. For example, with the well-known USAF 1951 resolution target, which is well accepted, well established, we can get some measure of the modulation or contrast transfer function, and that can even be done in multiple focal planes if you wanted to do some additional analysis of the depth of field and the axial performance of the system.

Also, this is the type of a target available from Edmond Optics, which is simply a very precise grid pattern which could give us access to characterizing the distortion produced by the system as well as the scanning accuracy, that is the alignment of multiple images that the system acquires.

These are just a couple of examples of artificial targets, but we could also imagine using a biological specimen as a test target during physical system testing by basically performing some sort of quantitative analysis of a specimen, and this is an example from a recent paper by Dr. Klaus Kayser, et al., and basically what you have here is an H&E stained slide, and what we've

computed here is the histogram of the intensity in this image, that is the brightness across the image, and that kind of quantitative metric could then be used to validate device performance and also track device performance over time.

So as we think about how we're doing this testing, we need to make sure that we capture certain pieces of information with each of these measured characteristics. Certainly we have to determine and declare the test target and the procedure used with that target. As I mentioned, we need some sort of quantitative figure of merit with each measurement, and then a benchmark or specification, and in this case, the conventional optical microscope could serve as the benchmark here.

And also as I mentioned, since we're interested in postmarket quality control, establishing a schedule for QC will be very important particularly for certain components such as the light source and mechanical scanners which can degrade over time and therefore may require frequent QC checks or calibration.

We've also heard from Dr. Descour's and Dr. Badano's talks this morning about standards, and these will certainly be a resource for us in developing test methodology particularly for individual components. There isn't much in the realm of looking at the microscope as a whole except for one standard from the German Standards Institution, DIN, which talks about quality management in microbiology, and this could potentially be utilized for

looking how to make sure that the WSI systems for histopathology are also maintained properly.

So, in summary, really what we're left with is the overarching question for the Panel as well as for the FDA is how to establish the proper physical testing guidelines, and certainly the standards provide some testing methodology which is relevant here, and given that microscopy is a several hundred year old practice, there could be also other resources and literature or from textbooks that could be applied here.

And, along with this, we would like to establish which physical tests are the most useful to validate device performance without excessive burden to the sponsor. It would seem that system level tests are the most straightforward, but also doing component level testing rigorously also can permit easy replacement of components as system configurations evolve, and that's a lesson learned from the radiology world, which we'll be hearing more about tomorrow as well. And as we think about these physical testing guidelines, we should make sure to keep in mind the variability within and between systems that can be tested very effectively with these procedures.

So with that, I'd like to hand it over to Dr. Robinowitz, the Pathologist and Senior Medical Officer in OIVD who will discuss the second category of preclinical and bench testing which involves the human observer.

DR. ROBINOWITZ: I've been a Medical Officer at the FDA involved with the regulation of in vitro diagnostic tests since 1990 and sitting

here, a couple of thoughts came to my mind and one is, there's been an evolution in the kinds of products we review, but also there's been an evolution in the way the FDA regulates these processes with more and more team reviews as you see today from the mix of speakers today and tomorrow from various parts of just our center. And so we're moving with the rest of medicine and science to a more systems-based approach.

So what I'm going to be presenting is some possible phantoms that might be used to test the human observers, the pathologists, with this new technology.

And I'd like to begin with this simple exercise in object-based diagnosis, and as Dr. Balis pointed out, this projector is suboptimal, and even my laptop is suboptimal, but what I wanted you to do was to look at this image and find the various objects in this image, identify them, then interpret what you think they represent, and sum it all up to a diagnosis.

These were taken from a paper published by Klaus Kayser and his colleagues in Germany, and what they wanted the reader to do was basically to replicate what a surgical pathologist does. A surgical pathologist detects objects and features in an image, determines quantitative and qualitative features. In this case, the little insects' size and feet and classify the detected objects, and if you notice, if you can see in the front rows, these little insects are quite musical. They have

instruments, and there's a little figure down here with a baton, and the

summation of all this information was the diagnosis that this is a music band, and it's very interesting that they use the term ants for insects, but my colleagues point out, and I'm convinced now, that they're crickets. So we looked up to make sure the German language doesn't use the same word for ants and crickets, and indeed just like in English, they're totally separate words. So I don't know where the notation, whether there was an error in translation, an error in transcription, but would we call ants a wrong diagnosis? Would we call insects a correct diagnosis and the most correct diagnosis crickets?

So how can we objectively evaluate image quality as independent as possible of the possible subjectivity, the nomenclature, in exactness and so forth?

So we've been debating how can we quantify the contribution of image quality per se to the final tissue-based diagnosis that goes on the patient's chart? How can we quantify the contribution added by the eye and brain of the pathologist; also, the training and experience of the pathologist, and just a basic skill level of the pathologist? How can we quantify the ability not just to interpret the findings, but the search and detect function of different objects rendered by the optical microscope, OM, and whole slide imaging, WSI? How can we identify the contributions of ergonomics and the human factors of different workstations?

How can we use preclinical bench testing to the best

capabilities? We heard my colleague, Anant Agrawal, talk about physical phantoms. So we've been trying to think about biologic phantoms that also could provide objective data for prediction of the diagnostic performance of pathologists with the optical microscope and digital whole slide imaging systems in real world situations in a manner similar to what x-ray phantoms have provided to the radiology community.

The general application of WSI, as pointed out before, to routine surgical pathology practice potentially involves all of clinical medicine conditions, any organ in the body, neonates, actually fetuses to the oldest of the old.

All clinically essential areas within the whole spectrum of surgical pathology, even if we wanted to, may not have to be tested. They cannot be tested realistically. So how can we be smart about this in assuring safety and effectiveness, which is what Congress wants us to do, what the public wants us to do, what my mother wants us to do.

So human observer studies with biologic and human histologic specimens can serve as phantoms to minimize the subjectivity of study pathologists by emphasizing each pathologist's assessment of image quality and information content before the overlay of high level, intellectual diagnostic interpretation by the eye and brain, the brain of this physician pathologist who's spent a lot of time learning all the features to look at, the rules for diagnosis and so forth, but how well does that individual actually

see? How good is that image? Is it a suboptimal image like I displayed in our little exercise, or when is the image good enough?

The targets can be constructed with biological materials of known size, number, fine morphological features, and spatial distributions to quantify image quality objectively. That's the key to what we think would be an ideal biologic phantom. It is controlling for the subjectivity, identifying the subjectivity from the pathologist as a human observer.

So every step of the line of the testing, whether it's bench testing or clinical testing, we feel should have a pretest tutorial to standardize the criteria, the nomenclature, as one of the public speakers mentioned, for the study pathologists. Everybody should call the insects ants or crickets or butterflies, whatever the terminology that's desired.

Then pre-specify the features to observe, enumerate, using criteria that minimize requirements for higher-level interpretation by the pathologists, analogous to the Snell eye chart. We're testing image, how well do you see it? Is it being delivered?

The task is recognition of shapes, sizes, images using descriptions rather than diagnostic terms and then to try to have the pathologist quantify as well as qualify, if that's a word, detailed features.

And I have found that there are diatom preparations available commercially, and just as a point mentioned by Dr. Rosai and Dr. Netto, we got this slide from a firm in the UK, and it took 13 days by airmail to get this

slide because it had to go through U.S. Customs because it was biologic.

On this slide, we see, in this photomicrograph, a section of a glass slide that has 97 different diatoms that range, we estimate, from about 2 to 100 micrometers in diameter. Despite being so small, biology is amazing, you can see internal 3-dimensional structures that can be objectively measured and enumerated to an extent beyond what physical phantoms allow us to do. At least we weren't able to find any physical phantoms that can be used at magnifications of a 40X microscope objective.

So we think that perhaps some clever people might be able to identify other biologic phantoms. This slide cost $28. It's really cheap, and you can get thousands of them if you want. Diatoms are the algae in plankton. So it's really inexpensive.

Whether it's diatoms or human specimens, we at the FDA have been thinking that an objective open book exam, at least for part of the bench testing, would help control the subjectivity of the pathologist so that we could see the results, and we are results driven in our regulation, of non-subjective evaluation of image quality rendered by the optical microscope versus the image quality presented to the human observer rendered by WSI systems.

The goal is to have reproducible optimum preparations of normal tissue and abnormal human specimens. For some purposes, tissue microarray (TMA) cores would be good, but for some, whole slide images will be necessary.

So what would we expect the human observer to do? One of the things that's happened in the last 20 years is that we're going more and more to a common standard for in vitro diagnostic tests, and all diagnostic tests, for that level of accuracy and precision that should be measured for each type of diagnostic test, whether it's a cardiovascular test or an in vitro laboratory test, whatever test, there should be a measurement for accuracy comparison to the best approximation of truth. Some people call that a gold standard. Most of the time, we have imperfect gold standards, but whatever the best available diagnostic truth is right now, that comparison is accurate. Then we want to know how the variability, the random variability of the pathologist looking at these objects can be measured when looking for features within these target objects.

We'd want to make sure that clinically significant objects were subjected to testing, and we just came up with a few that Dr. Becich shared with you, cancer microinvasion, nuclear chromatin structure, nucleoli, cytoplasmic vacuoles of various types, inclusions, granules, striations, architectural growth patterns, etc., are all the kinds of things that surgical pathologists must be able to find in their examination of human specimens.

With bench testing, we think it would be an efficient way to vary the effects of tissue preparation, slide thickness, different stains, different adjustments in settings of the optical microscope and whole slide imaging systems, variability in the planes of focus, selected by the WSI on

different runs, variability in pathologists with different types of image, allowing a large number of pathologists being able to see the very same slides.

The bench test could mimic what the surgical pathologist does in routine practice. So task one of a pathologist is to search and find and detect objects and textures that have to be examined and would require image quality, focus, fidelity, dynamic range, color fidelity across the whole slide, not just pieces of slide, but every single bit of, every morsel of human tissue of the slide. There are tasks for enumerating the number and measurements of objects, but you have to be able to see it well before you can count it and measure it. And then after you've done that, you classify what these objects are, and these three tasks could be done with cores.

There are tasks for enumerating the number and measurements of objects, but you have to be able to see it well before you can count it and measure it.

And then after you've done that, you classify what these objects are, and these three tasks could be done with cores.

I put a black box FDA warning around test 4 to emphasize that the diagnosis must be tested with whole slide specimens because the diagnosis depends on looking at everything on the slide and being able to enumerate and score the objects and features according to whatever the rules-based diagnostic scheme is being used. A particularly difficult test we

think would be quantifying what the pathologist is seeing, possibly with some scoring system, and whether optical resolution at 40X and Z-axis focusing would be involved. These test procedures would be something that we think might be performed on appropriate biologic phantoms.

Let's look at the principles of tissue microarrays. They are a very efficient way to display multiple cores, multiple samples from multiple paraffin blocks and assemble them within one paraffin block so that they can be cut simultaneously to prepare a glass slide with multiple specimens. Everything depends on who the expert is who selects these features from the whole section. So that's number one requirement.

And number two is that when the sections are made, somebody has to ensure that the particular slice of those cores still shows the claimed feature. That said, tissue microarrays come in various sizes from .6 millimeter in diameter, 1 millimeter, 1.5, 2.0. The most popular is 1.0 millimeters, and to give it some metric here, .6 millimeters in diameter is the standard 40X objective field of view, and 1 millimeter is the standard 20X objective field of view. Larger cores can be made, but even 1.5 millimeter and 2.0 millimeters are very hard to work with, but they are good for demonstrating all the architecture and features in normal tissue if you need to do that.

A .6 millimeter in diameter core allows you to put 500 cores on one slide, a 1 millimeter, 200 cores per slide, and since 1 millimeter is the diameter of a 20X objective field of view, that means that a 20X objective

possibly would have 200 fields of view on a slide.

So tissue microarrays have some benefits, and they can be used for recognition of cellular objects and features, but they're limited in discerning spatial tissue growth patterns, and they're not suitable for the full testing of the search and find requirements of making a diagnosis in the real world.

Another possible use of bench testing might be in very realistic applications where you have a whole slide with various pathologic features, and before the pathologist is requested to examine the slide, the pathologist could be presented with a schematic of the types of objects that are to be looked for and the nomenclature for recording the observation. So it would be a relatively simple interpretation to know that this is ductile carcinoma in site 2 that you're going to be looking at and these are the four patterns. Which of the four patterns is present, and did the image present enough sufficient details for you to make the match and describe?

Another test might be something like the Nottingham breast cancer scoring system, an ordinal system where for a pathologist finding tubules, mitotic counts, and nuclear pleomorphism would have a minimum of subjective intellectual interpretation and overlay and could depend more on what the image quality is, and the enumeration would be a measure of what the quality of the image is, and this is just one of many clinically relevant types of possible platforms for bench testing.

Also from Dr. Klaus Kayser's group, Dr. Mireskandari published a possible hierarchy for how to evaluate diagnostic errors or discordant results according to a 10 level increasing complexity of diagnostic interpretation and final diagnosis nomenclature. So 0 is no information. 10 is the final diagnosis. 1 is at the level of organ tissue? Is it normal lung or abnormal lung? If abnormal, number 2, is it inflammatory or neoplastic and so forth. And all the way up to whether it's a low grade cancer or a high grade cancer, I believe that's pretty much image based, but when you get into subjects like prognosis, good/fair/poor, therapy follow-up, no need/yes, that a pathologist does render in the diagnosis, that really has to have high order intellectual overlays. So I think the bench testing would focus on the image based parameters, and a clinical pivotal trial would bring in these interpretations. Does the pathologist come to the same identical diagnosis?

So we're going to ask some formal questions in the question period, but just to show you the kinds of thinking that we had in planning, we want to know what roles bench testing of optical microscopy compared to WSI systems can serve in guiding the designs for clinical studies of real world surgical pathology practices?

What safety and effectiveness concerns can be resolved in the preclinical studies using selected histopathology specimens such as TMAs and whole slides? And what are the limitations?

Finally, what concerns require clinical trials with

representatives of the intended pathologists population, the Juan Rosais all the way down to the first year resident? And who must examine and when do you need to examine the whole slide in the real world setting of surgical pathology practice?

Thank you for your attention.

DR. ADCOCK: Thank you. At this time, I'd like to ask the Panel members if they have any questions of today's FDA presenters. And this includes the presenters in the last part of the morning as well. Yes.

DR. SINARD: Hi. I have a question actually going back to our first speaker, Ms. Faison. And I don't know if you're eligible to be asked, but I still have a question. You know, I think that there was, as I understand from your presentation, the microscope is considered a Class I device, and you made the comment that whole slide imaging cannot be considered a Class I device, and I'm not sure I understood the reasoning behind that.

MS. FAISON: I'm sorry. I didn't hear the end of your question. He was whispering in my ear.

DR. SINARD: No, no problem. If the microscope is a Class I device, and I guess my question is, you made the comment that whole slide imaging cannot be considered a Class I device.

MS. FAISON: Correct.

DR. SINARD: And I'm wondering why? What is it that you do with whole slide imaging that you don't do with a microscope?

MS. FAISON: It's not what we do that's different. It's the technology. I did make a comment about Class I devices, and the second slide that I showed that showed the limitations to exemptions for microscopes says very clearly that if there's a new technology for that intended use or if it's an in vitro diagnostic used for diagnosis, then it's no longer eligible for that exemption. And that would, what we call, trip the limitation of exemption, and therefore it would be subject to premarket requirements.

DR. SINARD: Okay. So --

MS. FAISON: Because something is Class I does not automatically make it Class I. When new technology comes along, then naturally it would, you know, it just can't continue to be Class I.

DR. SINARD: Okay. So I guess my, and I apologize if I'm a little thick here, so are you saying that if the microscope didn't exist before and would come along today, it couldn't be a Class I device. It's Class I because it just -- it predated the FDA essentially.

MS. FAISON: Right. Yeah, can you help me with that, Bob?

DR. DECKER: Okay. So regardless of the means by which microscope actually reached a Class I exempt status, there are limitations that apply to that exemption, so that the citation would be in, I think, 866.9 which includes in it several subparts, one of which is a new technology that can be applied perhaps for the very same use as the original exempt apparatus would have been applied to. In the reg, actually what they cite as an example

is a scalpel versus a laser that's used for cutting. Both might be used for the very same kind of activity, yet one could expect that there are new questions of safety and effectiveness that might attach to the laser in distinction perhaps to the scalpel.

And so the reason goes that in this case, you're looking at a technological advancement or technological modification or one expects an improvement that trips that limitation to the exemption.

DR. ADCOCK: Dr. Hewitt?

DR. HEWITT: Directly related to that question, and that is the addition of changes in the light source for a microscope, that does not trip that limitation, the introduction of LED light sources or light pipes on standard microscopes as opposed to the traditional filament-based illumination?

DR. BECKER: I guess that's a hypothetical that we haven't actually needed to consider here.

MS. FAISON: It depends on the intended use. So, yeah, it's hypothetical.

DR. HEWITT: I mean they are commercially available with LED light sources. I can obtain an LED light source for a clinical microscope from a number of vendors today.

DR. BECKER: Possibly so, but that's actually outside the scope of our discussion here today.

DR. ADCOCK: Dr. Gilbertson.

DR. GILBERTSON: You talked a lot about -- things and seeing the details of things as pathologists and -- have to do that?

DR. ROBINOWITZ: I'm sorry.

DR. GILBERTSON: You talked a lot about a series of tests you would do. One of those was to illuminate -- allow the pathologist to basically count or measure things. I'm a pathologist. I don't do that.

DR. ROBINOWITZ: Well, I was presenting some ideas that we had. These are not regulatory requirements. But the idea of enumeration, counting or measuring was because we thought that if you enumerate, you have to really be able to see it, and pathologists as you know well look at patterns, and sometimes it may be a blurry image, but you know from experience that that is, you know, something that you've seen before very sharply displayed, but you know it's that.

DR. GILBERTSON: Okay.

DR. ROBINOWITZ: So it was really not a clinical practice issue but rather characterizing images as they're presented by various optical microscopes and various whole slide imaging. And I think one of the beauties of this Panel meeting might, at least for me, is that characterizing what the light microscope is actually doing and the range of it and the lack of specifications is very important because you can't go there if you don't know where you're starting from.

DR. GILBERTSON: Okay. I appreciate it. I just --

DR. ROBINOWITZ: Thank you.

DR. GILBERTSON: I think that's fine, but I think -- it's okay. I just -- what I didn't see in either of the presentations was, is the issue of when things are not there. I mean the one thing that scares me, I'm not at all afraid of giving a pathologist a virtual microscope and go after it because they're going to know if they see something that's blurry, oh, it's blurry, they'll deal with it, I think. Maybe not, but I think they will, but what does scare me is when the robot who's imaging that slide, pathologists used to do that job, I would read the slide with the microscope at least when I was there, we're taking that away, and I think that's something -- I'd love for someone to think of that very carefully, how we'll measure that and make sure that is always done right. That's one place you'll have a problem.

DR. ROBINOWITZ: One of my slides had fit for purpose. That's a very popular term now in development of diagnostics, is it fit for purpose? And what you're sharing with us or what you think a fit for purpose device should be, to be able to do this task. So I think one of the things the Panel might be able to suggest to us are what are the tasks that are essential?

DR. MELLO-THOMS: The questions are for you as well. You spoke a lot about image quality without really providing a definition that you were using for image quality. From your final slides, it seemed to me that the definition you're using for image quality was how much diagnostic information could be acquired from the image. So is that the correct

interpretation of what you were saying?

DR. ROBINOWITZ: That's my understanding.

DR. MELLO-THOMS: Okay. Now, based upon that, towards the end, you showed that a scale from a paper is that these features are imaged based and these other features require higher level intellectual --

DR. ROBINOWITZ: Actually that was my overlap.

DR. MELLO-THOMS: Right, right.

DR. ROBINOWITZ: The author didn't say --

DR. MELLO-THOMS: No, no, right. That was your interpretation. But for some of those differentiations, like when you're trying to differentiate between sarcoma and carcinoma, wouldn't you say that that's requiring higher level intellectual? I mean it doesn't come from the image. It comes from how you perceive the features being displayed in the image. So I don't understand how it's ever possible to separate the intellectual properties of the pathologists from the features in the diagnosis being rendered from the image.

DR. ROBINOWITZ: Well, you know, I think anatomic pathology has been characterized as sort of a right brain activity, very good artistic interpretation. Pathologists remember images that they saw 10 years ago, and this sort of stuff, and I think the reason we have so many different experts from different disciplines is to deepen our understanding of what the pathologist does so well. I mean nobody reads their own slides. They always

have a pathologist read it, and so whatever that pathologist is doing with the eye and brain and stuff, we're trying to quantify now, we're trying to understand and hopefully make it even better. So that was my -- I was just trying to see how can we objectively interpret how the human eye and brain are seeing that image based on what the instrument is delivering to us. And I'm not an expert in this field, and I'm grateful that you're here as an expert in image interpretation to help guide us. Is that a --

DR. MELLO-THOMS: Well, I have a lot of questions for you, but we can talk afterwards.

DR. ROBINOWITZ: Great. Super. Actually what we hope is that what we learn from here will help serve for a guidance on this for FDA and for industry. So if you have input to send to us after the Panel, that would be greatly appreciated.

MR. BRACCO: Dr. Robinowitz, are you responsible for the first presentation as well? Are you speaking for that?

DR. ROBINOWITZ: No, sir.

MR. BRACCO: Okay.

DR. ROBINOWITZ: I think Mr. Agrawal.

MR. BRACCO: Thank you.

DR. ROBINOWITZ: He's the engineer.

MR. BRACCO: Okay. Good. One of the slides you presented, you listed some imaging performance characteristics to consider, for FDA to

consider as preclinical testing, and you listed those items at the component level. And as the Industry Representative, I'm very concerned that some of these tests may be more academic and may not have any merit towards the clearance or approval of the device.

I'm also concerned that established pass/fail/acceptable criteria for some of these components may be unattainable.

So my question to you is how did you derive this list, and what are we trying to communicate to FDA here? This is a menu that you could choose from, or do you believe that these are the most important components that we should be looking at at a preclinical level?

DR. AGRAWAL: What I think I was trying to convey was simply first of all just really make sure everyone is in appreciation for the complexity of these devices. I don't think we had looked at it up to this point during the day, looking at all the components involved in these systems. So really I think to answer your question directly, it was meant to be a menu, not a list of these are the required items that need to evaluated at the component level, and at the end, as I said, one of the questions for our esteemed Panel here is to help us determine which of those tests, the component level, but perhaps mostly at the system level, because again each of those components such as, you know, looking at the resolution and contrast, I mean there is multiple components, each of which contribute their own effects on the resolution and contrast in the final image. So perhaps looking, just looking at the sum total

could be sufficient in some regard, but really what I was trying to present is simply a menu of items that we could think about, and there are more items on that menu that I didn't present, and we heard some of those from Dr. Descour, from Dr. Badano, and we may hear more from other speakers tomorrow as well.

So really I think it's just really trying to get us brainstorming about what is out there and then what's important.

MR. BRACCO: Okay. Thank you.

DR. ADCOCK: Dr. Zhou.

DR. ZHOU: Yes, my questions was on the overall purpose of doing the preclinical bench testing. If you recall, there's a different definition of accuracy of diagnostic test, depend on which way you looking at. So one way to define the diagnostic accuracy is use high -- structures. So I think from the 1 to 6. So the first level is feasible features you talk about here. The sixth level is diagnostic accuracy or distinguish between disease and non-disease.

I think we might want to think about when you looking at preclinical bench testings, maybe one should also consider the sixth level of accuracy, which is diagnostic accuracy. So to give an example, you could have very good imaging, but image itself does not tell us distinguish between disease and non-disease patient. So those kind of imaging is not clinical useful. So I think somehow we looking at the preclinical, we also need to take the second level of diagnostic actually into consideration where you try to

choose optimum features.

DR. AGRAWAL: Thank you.

DR. FORAN: This is for Dr. Robinowitz. I think the idea of using the TMAs is a great idea for looking for different features, but I know how much trouble goes into assembling these things, and I wonder if it would make more sense to actually create a virtual TMA, do microdissection on a whole slide rather than actually physically creating a tissue microarray. So in other words, using the same region of interest for WSI as for OM, without actually creating and constructing a tissue microarray.

DR. ROBINOWITZ: We're open to advice. I think what we were thinking of with tissue microarrays was being able to assemble let's say 100 samples per slide and making three slides from that paraffin block for three different readers, just a way to efficiently present lots and lots of features in this preclinical testing, and maybe to have each of the three slides stained differently. That was the idea. It was not in any way to try to mimic the surgical pathology diagnostic process.

DR. FORAN: No, I'm just suggesting that it might facilitate doing these high performance or high throughput tests. You wouldn't actually have to create the physical array. You could do it digitally.

DR. ROBINOWITZ: And the pathologist would just look at that one. Well, one of the problems --

DR. FORAN: We were thinking of comparing an optical

microscope to the same material shown to the WSI.

DR. ROBINOWITZ: Right.

DR. FORAN: So would that suggestion be able to work on the optical microscope?

DR. ROBINOWITZ: Yes, you would look at the same region of interest for WSI and for OM, capture both of those, and then present it to the pathologist.

DR. FORAN: Using the optical microscope?

DR. ROBINOWITZ: Yes. Just looking at the same region of interest.

DR. FORAN: With a motorized stage.

DR. ROBINOWITZ: Correct.

DR. FORAN: Okay.

DR. ADCOCK: Dr. Kulesza.

DR. KULESZA: I just wanted to -- maybe it's a little bit off topic, but through harping on the TMA idea, if you indeed want to do this the way by directly having a glass slide with 100 cores on a TMA and then projecting that using an imager, I think that that will lead to just practically horrendous problems because to score these TMAs on a glass stage and not lose track is almost impossible. So the practical issue of analysis of TMAs, and I know from painful experience is it A16, no, D -- gone, and I will have to go and count the cores again to make sure that I am actually looking at core 15C among the 100

that you gave me to analyze, whereas in a W, I can draw a grid. So, in essence, the wholly digitized slide will be error free because A6 is on a grid, whereas when I have to manipulate the stage, particularly under 40X where I lose track and cannot step back to 2X and immediately find where I am, even flipping between 2X and 40X in those situations is quite hard, and I lose track.

So from a practical standpoint, and I don't know if this is germane, I probably am using up too much time, but the TMA idea just for testing purposes, to trust a pathologist to score the same core would be problematic.

DR. ROBINOWITZ: Dr. Foran had a comment.

DR. FORAN: Yeah, I think that there are mapping software programs that can be utilized to help you navigate through a tissue microarray and make certain that you're returning to the same location.

DR. ROBINOWITZ: I think the key that we were thinking about for preclinical was to try to have it as objective as possible, and if that is a concern, to adopt some sort of motorized stage in the software to just present the images to the pathologists. It's not to replicate what you do in practice. Thank you.

DR. MELLO-THOMS: My question is to actually Dr. Badano who presented in the morning. You said that currently there is no standard way of measuring -- of color displays. And my question to you is how far away are we from such a standardization? I mean is it something that is going to

happen within the next year or in the next like 10 years?

DR. BADANO: That's a good question. I know of two efforts that have just started weeks ago. One is the extension of the TG18 document into color with the creation of another task group at APM. Those task groups tend to -- of course, I'm the co-chair of that task group, and we have the intention of doing this as fast as we can. Previous experience tells me that it would be very difficult to have any sort of draft before one year. So I think two years may be a timeline for APM to produce any significant work and publish.

The other effort is in the IEC, which is the International Electrotechnical Commission. There it's even worse because the level of discussions to reach consensus between manufacturers and organizations and government organizations and academicians across different continents makes it a little bit more difficult. So at least two, three, four years. The previous IEC I was involved actually took seven years to publish.

Those are the two efforts that I am familiar with and I can comment on.

DR. MELLO-THOMS: Do you see that as being a problem for whole slide imaging considering that properly displaying the image is an internal part of the process?

DR. BADANO: I think it's an important component. Let me just clarify. The standardization efforts are mostly geared to determine what

measurement methodologies you might want to use. Those will not solve all your problems. That's just the beginning. Once you have the methods, then you can start implementing those in software applications for quality control, like we've done in radiology. But again that's not the answer completely, the complete answer. So it's a long process. I think APM and IEC are committed to making it as fast as possible, and perhaps we talked about doing demonstration sessions at major conferences where we can gather support from industry and others to develop consensus in that area faster.

DR. ADCOCK: Dr. Birdsong.

DR. BIRDSONG: This is for Dr. Robinowitz. Perhaps you can expand on, earlier in your talk you had said that you had looked at developing phantoms and hadn't found anything that really worked well, and if I remember correctly, that's when you went to the TMAs.

DR. ROBINOWITZ: These are all exploratory ideas, and the fact that it's now a public issue, I'm sure some clever people will give us some good ideas, at least I hope.

DR. BIRDSONG: All right. Because I had some thoughts similar to the previous question about, you know, just reading TMA with 100 slides, I hadn't thought of a motorized stage, but I think really what we want to quantify is within some broad limits the color accuracy. I'm going to come back to that in a minute, and, you know, like I asked the question this morning, I think the dynamic range issue is really important, and maybe that's

because I'm, you know, a cytopathologist, and we've already established or stated that these wouldn't be used, but even with histopathology, I still think the dynamic range issue is as or more important. As Dr. Balis pointed out this morning, some of the best, you know, diagnosticians on the planet have, you know, some problems with color perception. So, you know, I think it's easy to make a little bit too much out of absolute color accuracy. And one of the questions I think you want to ask as you investigate it is how accurate must the color be? I mean obviously something that should be orange or pink can't come out blue, but if it's one or two shades off, is that really going to make much of a difference?

DR. ROBINOWITZ: Well, I've read that pathologists look at controlled artifact. Tissues are colorless until we put 150 year old stains on it and it becomes hematoxylin and eosin, and I think if you are colorblind, whatever that is, I'm not colorblind, but if I were colorblind, I would know what is, and I was trained to look at many, many images, the truth for me would be shades of gray and purple and pink for you.

DR. BIRDSONG: Right, and that's why I say the dynamic range is important because that's what I think the color --

DR. ROBINOWITZ: Absolutely.

DR. BIRDSONG: -- are seeing.

DR. ROBINOWITZ: I think Dr. Balis pointed that out very well.

DR. BIRDSONG: So I think one of the things, one of the

questions you want to ask as you investigate this is, you know, what are the boundaries, if you will, of what we need, in terms of adequate color accuracy, and similarly for dynamic range, he also pointed out that the dynamic range of current systems is substantially less than what you might actually see in a specimen, yet the studies of diagnostic accuracy have shown that, you know, current systems are pretty good. So at what point does a system not demonstrate sufficient dynamic range to be sufficiently accurate for diagnosis?

DR. ROBINOWITZ: Well, I think the point of the preclinical testing of biologic specimens would give you a way to have a controlled study with limitations of interpretation of the image difference with a slide that's let's say a lymphoma with different thicknesses, different stains, different devices, different light source, whatever. I mean it's just a platform for answering these, but from the clinical pathology world, these would be analytic studies. In the analytic studies, you want the limit of detection, you want the limit of quantitation, precision, all that sort of stuff. So we were just thinking at how could we apply what we do for clinical laboratory tests, blood tests, immunology tests, and make the evaluation include pre-analytical issues, analytical issues, and post-analytical issues.

DR. BIRDSONG: Yeah, I'm in complete agreement with that. I was just saying there seems like there's so much inherent variation in a real biological specimen that if you could break it down into more specific, you

know, parameters in the terms that we've already discussed, and if a system meets, you know, this degree of color accuracy and had such and such dynamic range, it may come down to how many bits you're digitized at, but can it accurately transmit that? Do you need a certain level of ambient light to perceive that accurately, and questions like that. You could probably come up with a set of parameters that defines a system with adequate performance, and you would obviously want to correlate that with biological specimens, but you might -- the bottom line of what I'm trying to get at is we -- I think you want to try at least to get around some of the inherent variability that comes along with using biological specimens, to be able to spec out a system that will work adequately with biological specimens and you won't have unit-to-unit, you know, variability that you can't separate from the inherent variability of the biological specimen.

DR. ROBINOWITZ: Your question sort of touches on Dr. Sinard's question, and that is in the risk-based regulation of the FDA, the more you know, the more you can mitigate safety issues. For example, a potassium level, if it's off, can kill a patient, but we have robust, accurate, precise measurements to apply to potassium. So the risk is mitigated. But if you don't know all the factors that may have a critical influence on the new test system, then there's an inherent risk from lack of knowledge. So the preclinical testing could address these very important issues that you're guiding us with, that you have to look at everything you know and see if it

really influences the image quality to the human observer.

MR. BRACCO: I think it's important, though, that we differentiate between what has an impact on the preclinical study and what has an impact clinically because we are going to find differences, and it's very important, and I'm nervous that we're on a slippery slope here, that we're going to ask for preclinical studies that really have no merit towards the final device and its intended use.

DR. ROBINOWITZ: Well, actually, I think what we're doing is really a general set of concerns and a general set of recommendations for anybody who is designing and manufacturing a whole slide image. It's very much like so many things in industry where there are industry requirements that every manufacturer has, even before they put a label on their product. And also Dr. Birdsong, Dr. Davey, Dr. Kulesza, Dr. Gilbertson, Dr. Sinard, Dr. Hewitt, all of the pathologists who would be considering these devices would want to know how can I be an informed consumer. What really matters? What doesn't matter? And I don't think it's singling out a particular manufacturer. I think all the manufacturers will need this because the whole quality systems design requires knowing what the purpose is and fitting that purpose with the final product.

So I think what we're talking about today is trying to know -- we have a pretty good idea what pathologists need, but we're not sure how those needs can be quantified and separated and interpreted.

MR. BRACCO: I think being informed, though, is quite different than collecting data to be used as part of the device's clearance. I just want to make sure that we all understand that there are things that are informative and academic, but there are also things that we're going to need to evaluate these devices to get them cleared, and it's very important that we distinguish between the two because it would be unfortunate for manufacturers to have to run through a whole gambit of tests which FDA then has to analyze and figure out what to do with when they really have no merit on the clinical performance of that device. So that's the point I'm trying to make here.

DR. ROBINOWITZ: Excellent point.

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: And so following up on that question, am I correct to interpret that there is a concern or we should consider the issue that the instrument now interceding the place of a Class I instrument, the microscope, is true in its reproduction? Hence, we would be concerned that each instrument would comparably reproduce the same color and the same resolutions. Is that a correct concept or interpretation of some of the concerns? I want to make sure that instrument X and instrument Y from either the same or different manufacturers, if I have three in my own lab, are all going to get back the same color just like I have the expectation that if I go from microscope to microscope to microscope to microscope, they would always turn the same color. Are those reasonable expectations?

DR. ROBINOWITZ: Well, as Mr. Bracco pointed out, the regulation of these instruments should be based on the clinical impact supported by data, and if people can demonstrate it doesn't matter, fine, but I think the uncertainty right now is we don't know, at least I don't know, just what the impact is of these variations that you were describing. Do you have some concerns?

DR. HEWITT: Well, I think that you all demonstrated some slides earlier that showed, you know, misadjustments in color, making interpretation impossible, and when you look at those through the microscope, it's one thing, but when you have an instrument that intercedes in your interpretation and you have not had an opportunity to review the other, you may not know that the object that you are viewing is not an accurate representation of a truth function. So some element of calibration I think is something that needs to be discussed.

DR. ROBINOWITZ: Well, calibration is basic 101 clinical pathology methods evaluation.

DR. ADCOCK: At this time, I'd like that we should take about a 15-minute break and resume at 3:40.

(Off the record.)

(On the record.)

DR. ADCOCK: At this time, we will begin our discussion of the FDA questions. A copy of the questions are in the Panelist folder. Ms. Faison

will read the questions, and we will project them. So please show the first question.

MS. FAISON: The FDA is asking the Panel for advice and recommendations for the scientific and technical characterization of the physical characteristics of images produced by the conventional light microscope and digital whole slide imaging systems -- according to the requirements to maintain the present level of accuracy and precision expected or required for safe and effective routine surgical pathology use.

Each of the questions about the optimal bench testing methods and studies needed for the objective analytical evaluation of the physical characteristics of optical microscopy and whole slide imaging may involve one or more of the following parts of the OM and WSI modalities:

- The optical physical features (magnification, numerical aperture, working distance, xyz-limiting resolution, wavelength band, field of view, depth of focus, stroke, focus resolution, ocular tilt range, optical axis offset, light source, color spectrum, etc.)

- The image acquisition process (the color and spatial resolution, sampling, bit depth, and sensitivity, as well as field of view, frame rate, shutter speed, quantum efficiency, gain control, output format, etc.).

- Image processing (compression, tile stitching, smoothing, sharpening, etc.)

- And the display system (display type, size, image resolution,

dynamic range, color calibration, etc.)

Question 1: What U.S. or global standards, guidelines, industry operating practices are available to guide the characterization of OM and WSI?

DR. ADCOCK: At this time, would any of the Panel members like to provide -- Dr. Kulesza.

DR. KULESZA: No.

DR. ADCOCK: No guidelines exist.

DR. KULESZA: No comment.

DR. ADCOCK: Dr. Gilbertson.

DR. GILBERTSON: I know of no useful standards in this area. I would say this, that the optical microscope as used in pathology has been evolving over 100 years, and there's a number of companies, all competing in this case, and all of their microscopes are relatively similar. So I think the only -- is there's been, you know, 100 years of pathologists buying systems, at least from the OM side, they've evolved and they're pretty standard.

DR. ADCOCK: Dr. Sinard?

DR. SINARD: I actually don't have a direct answer to the question, but I'm going to use this forum to just try to -- I mean I think it's important. My perception is I'm here to express my opinion, so I guess you're going to get it. I think it's important to keep this whole process in perspective and to revisit what I discussed before. The actual process that we're looking

at here is the patient, lesion in a patient to ultimate diagnosis. And there are an awful lot of steps in that process, and I just jotted down a few of them here. So there's the surgical sampling step. There's the preservation and fixation step. The selection of what to submit for histologic examination. How the tissue is processed. How the tissue is embedded. The microtomy, both for artifacts, thickness, number of levels, all those issues. Then there's the staining and the cover slipping process.

All of this before we even produce the slide, that then goes onto the microscope that creates an image in the pathologist's brain which they then interpret correctly or not.

And I think what we're talking about is, you know, obviously just looking at an alternate pathway for that last step in a very long process. Each of the steps has a tremendous amount of inherent variability and interestingly absolutely no regulatory control for any of those steps.

And so I think we just need to be cautious when we start to try to come up with standards for how we are going to decide whether or not an instrument is operating properly. I don't think that it is necessary for there to be a restriction that the digital image be a precise portrayal of all of the characteristics of what is on the glass slide because a glass slide, there's a tremendous amount of variability in our source there. And so I would think that this step and whatever criteria we ultimately recommend, the step should not be more stringent than the inherent variability in the system

anyway. And so I think that's where it's sort of going to be difficult in designing a set of criteria. It's important not to hold this one step to a standard level that none of the other steps even come within an order of magnitude of meeting along the way.

DR. ADCOCK: Dr. Gilbertson.

DR. GILBERTSON: I would second that. I would say more than that. I would say that in pathology, the image is a chemical process that happens on a glass slide in histology. That's where we create our image. Before we go to stain things, light goes right through the tissue. It just goes zip, right through. You don't see a thing.

The microscope magnifies this image and there's other interesting things and important things, but for the most part, the image is made in the histology laboratory. It's not like radiology at all. A very different process. It's back-end also.

In radiology, the radiologist is given a study to read. The pathologist is given a piece of tissue and says what is it? And our job is not to read it but to do the test necessary for us to say I know what this is, and by the way, it's okay to take the leg off, right. It's a very different thing than you see in other studies, more than a study, and the reading of a glass slide is a tiny part of what this is.

DR. ADCOCK: Dr. Davey.

DR. DAVEY: I guess I agree and disagree. I do think that we

don't have as much control over the virtual slide. So I do feel like the instruments need to give us whatever scanning needs to -- we need to feel comfortable that they're going to give us a similar result every time given the same imager.

So I'm not saying that we should have standards across instruments, and there should be some level of precision that is looked at. I think we do need to, as things get more automated, we do need to look a little bit more towards what clinical laboratories do with their instrumentation, and I would just want to feel comfortable that I would get the same kind of scan since I don't have the ability to adjust the microscope. So I guess I would argue that and, you know, your concerns about capturing all of the tissue, those are the kinds of things that we could look at and, you know, calibration and so forth.

I guess I would argue that we would like to -- I went through and I used the thin prep imager, and we had to use a specific stain for that, and initially it took some getting used to, but there were some advantages to having a little bit of standardization, I would say, in things like sectioning, cover slips, and stain quality.

So maybe this is the time we should push for that.

DR. GILBERTSON: I would agree with that, but I would simply say that we're doing this at my hospital, but basically it makes our AP looks like a

CP laboratory, you know, robots. The robots do everything nowadays really, if you allow them to be, and it's really good. It's better quality. It's a 10 year job, and I think that we don't have the whole process, you know, stop development in one area of the entire process.

DR. ADCOCK: Dr. Sinard.

DR. SINARD: Yeah, and I just want to be careful as we go down the line of analogies to CP because I mean I think that AP can learn a lot from CP because CP has had a lot of these processes worked out. But I think a big difference is that I can't look at a blood specimen that the machine has told me the potassium is 10, I can't look at it and say, no, that's not 10 because I'm relying on the machine in that case to do the interpretation.

In anatomic pathology, there's still at the end of the day a pathologist in that loop, and so there's a pathologist who's going to use their medical judgment and their understanding of the artifacts and the process involved to decide whether or not they have enough information to make that diagnosis. So I think that if we were discussing here automated quantitation, which, you know, a lot of the speakers have started to speak to, as a potential future goal, then my level of concern about reproducibility and precision goes up significantly, but to the extent to which, and my understanding the scope of this discussion is limited to just the H&E diagnosis, not any quantitative process, that I think a large part of the responsibility for that still resides in the pathologist's judgment as to whether or not the quality is good enough to

make the particular diagnosis that has to be made in that particular case.

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: I think Dr. Sinard's a segue to my comments, and they agree and disagree in some fashion. I view the use of whole slide imaging as much as a test. It's a test. You're introducing a platform or an assay, and everything that's incumbent up into production of the slide would be pre-analytic, and there are a number of pre-analytic variables that you may suffer the consequences of.

However, when you have an individual assay, you go through verification of the assay, which means testing to the precision and accuracy of every step of the assay, and then once you know you have an assay that is verified, you come back and look at the use or the utility of that assay and the concept of validation.

I would suggest this would be an appropriate means of breaking up this challenge. We have the interpretation off the screen, which I agree with Dr. Sinard, is complex. It's not like potassium, and I think those are different subjects that need different requirements, whether they're H&E or maybe something else, but there still needs to be a calibration issue as well as a verification process for the production of the image because we are no longer dealing with just an optical microscope. I mean it's in some ways very much like going from the way my great, great grandfather diagnosed diabetes, which was sticking his finger in the urine, versus measuring urine

glucose.

And so, you know, you can break up that assay into the automation steps, the optical stops, the software steps and the display steps, and each can undergo a specific verification regime using different components of existing technologies, ISO standards and such, and so I think that although it sounds complicated, in fact, most of those technologies at least as a first basis exists, number one, and number two, they allow you to inwardly improve the technologies as they improve forward, focusing on the life cycle type of approach and trying to allow as continuous improvement of the technologies so that as new modalities become available, 40X scanning at Z and 100X, that we already have those things geared up and ready to go.

DR. ADCOCK: Thank you. Dr. Kulesza.

DR. KULESZA: I just wanted to better understand the comments that Dr. Gilbertson and Sinard made because it seems to me that we can compensate for a lot of the artifacts that are generated pre-analytically using the microscope because we can focus and we can look.

Now, if we scan at 40X and we generate a digitized image that does not allow us this ability, then what is your comment? I mean because that would to me mean that this box, this new machine, has to be extremely good, not to present me with a miss that I will not be able to interpret as an artifact to which now, as a predicate, divide the microscope, I am really familiar with.

So what is it, what are the requirements that you think should not be imposed on industry as unduly burdensome versus the ones that will actually make absolutely sure that the full scanning has adequate depth of field and I am not missing that cancer cell that's above the plane. Am I making clear?

DR. ADCOCK: Dr. Gilbertson.

DR. GILBERTSON: You're right. I agree with that almost entirely. I would agree I think with what you're saying in many ways, I think, because Mike Becich said something earlier today which is important, that pathology depends a lot on the pathologist understanding when he's seen enough because we can always go to electronic microscopy if we wanted high resolution. We could do it, right, as long as we know we have to do it.

So I would say that one thing -- I think that one of the things I think is worth looking carefully at is what are the steps in the process, the use cases, in which a pathologist, something might happen and the pathologist not know that it happened. They don't know they're being fooled, and if that happens, it's a real problem because we just lost our ability to control our process, and we know what those are.

Now, I think that, yes, if for some reason I'm looking at a slide and let us say -- and I see all of my nuclei are black, right, and I'm not seeing. My job is twofold. Either ask the scanner to re-scan the higher intensity of light or cut my slide thinner, right. It's my job to see what's there. My job is

to know what is there and know what is appropriate for the physician to do, right.

So that's true of all things. I don't think my machine has to be perfect for me. None of my machines are, right, even my electron microscope. I have to do another stain and this and that, right, but I think it's most important for this committee, I think, to identify where in the process could things happen which a pathologist does not realize something got lost, and that something -- I know of one place where it definitely could happen, which is when some of these devices -- they don't -- images in general, in general.

Industry will probably debate this and they'll get better also. There's no question about this. They never, ever design to be whole slide imagers because they want to be fast. So they say if you look at the slide at low resolution and say the tissues are there, I'll image the tissue, right, and if you fail and don't see all the tissue, you could basically cause a real problem. The pathologist might not know. You can design systems that makes the pathologist have a good chance of knowing. That's a separate issue, but I think it's something I think that -- it's one of those things where you could have, patients could be hurt by this technology, and the only time that could be happening is when the pathologist doesn't know something happened, and I'd like to know of in terms of studies where in this process could this happen, and are there solutions for this.

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: I think Dr. Gilbertson's comments match somewhat my own, and that is that the better the instrument is specified and the performance parameters of the instrumentation are understood, the better the likelihood of our capacity of assigning the errors in diagnosis, misapplication of the entire test, rather than the assay, which is the whole slide, to either the pre-analytic variables or the skills of the pathologist. What we want to know is that the machine is doing its function appropriately. In truth, you end up protecting the instrument by saying, if reported truthfully, and the failure was either on the pre-analytic human side or the interpretative human side because otherwise you have the risk of applying failure to an instrument for whom failure actually doesn't belong to the instrument itself but you have to specify that with great precision. It allows you to define where the lockouts are, where you shouldn't be. If the specimen's too thick, the machine should be, no, I don't want that.

DR. ADCOCK: Dr. Davey.

DR. DAVEY: Yeah, I think that echoes true. We don't want to restrict, you know, development, further development of these technologies. However, how studies are done need to be specified very clearly so that if you want to use the same instrument, you know how was the tissue processed, how was it sectioned, what kind of, you know, was it glass cover slips used, and otherwise, you know, you won't be able to -- laboratories won't feel

comfortable using the technique. I mean I don't see how you can say that they're directly comparable. So you need to be transparent in how things are done.

DR. ADCOCK: Dr. Zhou.

DR. ZHOU: Well, I agree. I think we have to have some standard to say that the instrument you're using actually reliable or reproducible. So if we don't have that, I think how can we trust whether the error you make is due to pathologist or actually due to the machine itself. So you first have to make sure machine is trustworthy. That's the first step, and that's what I talk about here. I assume you would have some standard for -- I mean not just say, well, leave to the pathologist. They can -- the results from the PSI.

DR. ADCOCK: Dr. Gilbertson. I guess I need to understand what you mean by truthful or correct. I mean for some reason, I don't understand. For example, I can see, I can take a slide, regular slide and I run it through a whole slide imager, and I find that my nuclei are underexposed in my lymphocytes. But in my macrophages, they're not. They're perfect. So is that a failure or not? I don't know. I mean I have a question. I have an opinion, okay, but, you know, it's up to me. In that case, I figure it's not. I'm talking about pathologist. My -- sees this and says, yeah, my lymphocytes are underexposed. I'll do something if I think it's important as opposed to another situation where, you know, the pathologist doesn't realize that

something bad happened. I just want to make a distinction in that area.

DR. ADCOCK: Dr. Kulesza.

DR. KULESZA: So it seems that there are two things here. One is that just like in CP, who do rheumatoid factors, etc., etc., interfering substances, matrix effect in -- or whatever we have, we should challenge the instruments with slides that are faulted, slides that are perhaps too thick, etc., etc., over-stained, and see what it does, whether the pathologist at the end will recognize those very common pitfalls, and we can name five of them in a guidance document, right. Chatter of the -- we all know this.

And the other thing is that from what Dr. Gilbertson says, it appears to me that what Dr. Robinowitz was driving at will become essential, i.e., how do we know that the instrument is good if his lymphocytes are perfect, but what's the truth here.

So we do need diatoms that are identical each time and perhaps a standard slide. Is that not the conclusion?

DR. GILBERTSON: No, because in this case, in this situation, we have two different cell types, right. One of them, you know, was very dense. It might need just to go through that, the thin light through that lymphocyte needs more light, right. So another question is do you image it twice? Well, maybe a pathologist might want to do that. It's possible. It's like another special stain, right, another setting in my device, I could -- twice, if I wanted to, but more likely if there's a case of, you know, a gallbladder removed for

stones, right. So I guess in this case a pathologist saw that the lymphocytes weren't perfect, but they were okay, right.

What scares me more is a situation where perhaps there was a whole, for example, the same case, but other things like this, there's a situation where a piece of tissue wasn't in, for example. In this case, I'll not know that necessarily. Maybe they will, but I won't, or something like that, but I think that can be used as a distinction between an image that is not perfect because I'm not clear that perfect images -- that perfect is not diagnostic. There's a difference there obviously. But if I'm a patient, I want to make sure that if a pathologist uses this newfangled device, at least I'm not going to put the patient at risk.

DR. ADCOCK: Dr. Sinard.

DR. SINARD: There's been a lot of talk about analytics and pre-analytics and the whole bit, and I think the suggestion seems to be that because these whole slide imagers are machines, that they are analytic machines, and I would actually argue that whole slide imaging is not an analytic step. It is part of the pre-analytic process. In an anatomic pathology interpretation, the analytic process wholly goes on inside the pathologist's brain.

I mean if you want to consider the whole slide imager as part of the analytic process and not part of the pre-analytic process, then you have to consider the entire histology process, you know, from the grossing in all the

way through the slide preparation, including the processing as also part of the analytic process because all those things obviously affect the ultimate interpretation of the presentation of the data to the interpretative point in the process. So I think that's an important distinction to make.

But I think the other thing that's sort of coming out of this discussion, equally important, is that so much of what is needed varies so much from case to case. And so to extend on Dr. Gilbertson's example, whether or not it's important to see the chromatin pattern of the nuclei of the lymphocytes depends on whether lymphoma is in your differential diagnosis, and that's a determination that's going to be made based on other morphologic criteria. And so what your differential is determines the specific thresholds for quality that that particular specimen needs to meet, and to hold another specimen to the same quality is potentially an extreme waste of time and effort and energy. Ultimately it is the pathologist who has to decide for this particular case, is the information I have enough to make the diagnosis?

DR. ADCOCK: Mr. Bracco.

MR. BRACCO: I am not a pathologist, but I wonder through all these conversations I'm hearing if some of these limitations aren't inherent on a regular microscope, and I'm wondering if we're holding the whole slide imaging device to an unfair bar here because I think we heard earlier today that a certain person went into a lab and determined that the microscopes

were inadequate, and he asked that they don't be used anymore, and I'm just wondering that if we're putting this new technology at an unfair -- raising the bar unfairly in that we're asking it to be perfect, when we've relied on microscopes for all these years knowing that they're not perfect. And all the manufacturers are trying to do is prove clinical equivalence and whatever the inherent deficiencies are in either device, clinically they're equivalent.

So again I don't know how well microscopes perform because I'm not a pathologist, but I assume they have errors where cells are missed as well depending on how they're set up, how they're stained, etc.

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: In reference to Dr. Sinard's comments, image making may or may not be analytic versus interpretative. Within the context of image analysis and immunohistochemistry, it would be considered analytic. I think that there is space for a great deal of restructuring of the way an assay consists. I've even proposed the term proto-analytic because staining in the H&E process would probably be considered analytic in and of itself.

I think, though, that gets back to the specification issues with regards to the lymphocytes, and the issue is, is the instrument capable of meeting the specification that is expected of it in reference to the gold standard, which is the microscope. And that gets back to your comments, Mr. Bracco, and that is, are we holding it to a higher standard? No, we're probably protecting it from a liability because if it is deemed unsafe, the

manufacturer is going to be the person who suffers the most, not somebody else. Right now the pathologist is assuming the liability that he is using an appropriate instrument to make his diagnosis, and that is kind of under the guild concept that was discussed earlier, and he assumes those liabilities, but you introduce a whole slide imager, you've introduced a new layer, and so a level of protection is incumbent, but it takes us back to specification. I think the more we know about the capacity of the instrument, the easier it is for us to apply a fit for purpose model of defining specifications for which the instrument is appropriate, and that may provide guidance towards us as to where it may be appropriate to design clinical trials or suggest that they are appropriate for diagnosis. It's unclear to me that an organ-based approach is appropriate because of fit for purpose.

Your chromatin, for example, where's it important, where's it not important. If we had a model in which we're working with a differential diagnosis, with interdifferential diagnoses in the benign realm, well, lymphocyte chromatin isn't that important, but all of a sudden as the concept of lymphoma popped into the pathologist's head, maybe you should say, hum, started thinking about lymphoma, I guess I need to flip this over to glass because that's the kind of guidance I see that we're wandering towards, but we have to define those specifications, and it doesn't seem to me that it's an easy thing unless we define what the instrument can do. I mean I don't think it's defined what the instruments can do yet.

DR. ADCOCK: Dr. Gilbertson.

DR. GILBERTSON: There's two points here. I'm not sure that we are asking for an equivalence between a whole slide imager and a optical microscope. I certainly am not because I think of cell phones, you know, cell phones are terrible. I mean a landline is so much better in audio quality, right. However, in my pocket and in yours, I bet, and yours and yours, there's a cell phone, okay. There are use cases where it's really useful and important, and I know if you're thinking about a whole slide imager, there's basically some digitization on top of a normal microscope. The question is why digitization? I mean, you know, is it for a lot of good reasons, not just image quality?

The other point though I think is the issue of a whole slide imager is doing something that is a little bit -- it's not a question of image quality here at all. Let's say they were completely equivalent. Let's say the whole slide imager was better, substantially better. There's still an issue which is that we're asking a robot to do a few things that the pathologist has always done, and I suspect knowing robots and knowing pathologists, we'll eventually create a robot that's just as good as a pathologist, even better. But I'm not quite sure we're there yet, and I think one of things that we have to make sure of on this Panel is that we don't do something or allow something to happen that might hurt a patient on the simplest things, you know, like finding tissue, whatever.

So I think there are two separate issues here, and I don't think the committee is really saying we want equivalence with the optical microscope. I don't think we do.

MR. BRACCO: If I could just respond. Maybe Dr. Becker might be able to help us here, but I was under the impression that we were looking for equivalence here and not that we had a better mousetrap, so to speak. Is that correct, that we're looking to see whether or not diagnostic equivalence has been obtained by whole slide imaging to microscopy?

DR. BECKER: Well, I think that's in part -- I'm sorry to hand this back to the committee, but in part that's a question for your deliberation. The questions posed are in the vein of being able to make a definitive diagnosis for the slide from the digitized image. So that certainly is the idea of being able to get at a diagnosis which is just as reliable as you would have been able to count on from the original optical microscopy is certainly within the scope of discussion.

Though what I think I'm hearing also is that you're considering some other venues, some other perspectives under which the device might be used, if I were to take, for example, the cell phone analogy that Dr. Gilbertson had brought up a short while ago.

DR. GILBERTSON: Yeah, I would say more on that. I mean for a pathologist to make a diagnosis, is not sitting in front of a monitor or a microscope like a radiologist and say, I see it. No, we're given a piece of

bloody, gory, miserable, you know, half fixed tissue, and we're asked the question, what is it? And our job is to do whatever test we can possibly -- every test we need to make sure we know what it is and that might mean doing something like this. Oh, yeah, this is all fat, right. I'll freeze it to death, and that's that. No microscope, right, yeah. Or it might be, you know, 50 -- stains, electromicroscopy, and review by an expert, right, but we do what's needed to get it done. So I think the idea of a diagnosis is being something that is, you know, we get that by looking at a slide. It's kind of somewhat weird. We don't do that. We have all kinds of tools including -- microscopes, including different objective lenses on the same light microscope, right, and we make those decisions all the time in our practice. I'll go to 40X. I'll go to 10X. I don't have to go to 40X. We do it all the time. It's our job. That's what they pay us to do, I think.

DR. BECKER: So I'm very attentive to not trying to guide or in some fashion add much to what the Panel's discussing here, but it seems to me that in part what you're debating here is the question of whether the device can be validated to have an appropriate place in the work process such that the overall process yields the right answer, whether to include information simply from the visual view of what the imager has given you, okay, in the context of other information versus whether you can, using the imager per se, get specifically to the right answer. It seems to me -- I hope I've made that clear. I want not to open a thread that is not what you're

actually considering, but it seems to me that that's the thread you are considering.

DR. GILBERTSON: I think that's what I was saying. Others may disagree, but I wouldn't expect -- I know that my optical microscope is never thought of, at least in my mind, that way. I mean it's another tool I use, and if I have to go someplace else I will, including FedEx is a way to get stuff around, but my job is to do what's needed, to get the right diagnosis and the right clinical information to the clinician, and it's a really serious job, right, because the clinicians are going to do some very nasty things to people depending on what I say sometimes.

DR. ADCOCK: Dr. Zhou.

DR. ZHOU: Well, I think we might mixed up with different level of evaluations. So -- about the design of the studies. So in diagnostic medicines I think is going to apply to here, too. There's three different level of the design, same as -- I don't think penalize for the diagnostic tool. Even for the -- development, there are phase 1, phase 2, phase 3. Similarly in the diagnostic medicine, we have phase 1, phase 2, phase 3, but we have different name.

So phase 1 is called exploratory design, and phase 2 is called challenge design, and phase 3 is clinical design. I think we hear actually the different type of criteria to evaluate them. They're different measurement. I think we are mixed up all three together. Here, if I understand correctly, we

talk about exploratory phase. Is that correct? Not talk about clinical phase. So clinical phase, the bar much higher to cross the bar, but for the exploratory phase, you have to show your diagnostic system can do very good.

So the way to select some point is you select very simple cases. Are they doing a good job with every simple case? Because it has diagnostic values. So I think we might be mixed up all three together.

DR. ADCOCK: Dr. Kulesza.

DR. KULESZA: I think that I agree philosophically with many of the things that have been said, but I just don't think that they are relevant at all. I think that with Dr. Bracco, I would answer this directly. I will not use the whole slide imager to replace any immunostainer or electron microscope. I will use it to replace the microscope. So it better be as good as a microscope because I will rely on it not to miss stuff.

So simply put, the only question in this realm, aside from the challenges, etc., and the specific guidelines for how to test the instrument, yes, I will consider it as a replacement of a microscope for the purpose of this evaluation because, of course, I will use FedEx and I will send you my consults, but that is not germane I think to the question at hand in terms of how to structure evaluation of a whole slide imager. Because the only tool that it's really replacing or conceivably is an optical microscope, no? No.

DR. ADCOCK: Dr. Sinard.

DR. SINARD: Well, I think it potentially could replace the

microscope for some cases, but the microscope is still there, and so it's not knocking the microscope out of the loop necessarily. And so I think that that's important to realize is that when warranted, based on the opinion of the pathologist, there's still the microscope to fall back on which has a different set of examining characteristics that allow you to do that. So I think that that's an important part of it. I don't think it is meant to be a complete replacement for the microscope. It can be in many cases, in many routine cases, but I think in a number of cases, one is still going to have to fall back on other analytic, if you want to use it, if we're going to use that term, or other examining devices, other tools, and one of those other tools still remains the microscope.

DR. ZHOU: Of course.

DR. SINARD: So in that sense, but I also wanted to speak to something Dr. Hewitt mentioned because you used the word liability, and I think that that's kind of a useful construct to examine the limits of what sort of level of performance we should be holding these various devices to. So I think that, let's take whole slide imaging out of the picture. What is the pathologist currently liable for? And how much of that can be attributed to the actual functioning of the microscope currently?

And I obviously, once you start throwing courts into things, the rules are not hard and fast, but for the most part, the pathologist is liable for not just what they look at through the microscope, and even not just what

happens to be on the slide, although at the very minimum, they're responsible for anything that happens to be on the slide, but arguably they're responsible for anything that was in the original specimen, and if it didn't end up on the slide, that's still the pathologist's responsibility.

So right now all of the liability of the whole process falls on the pathologist. It's an interesting concept that you introduced that if we insert an instrument into this process, can that instrument absorb some of the -- should that instrument absorb some of the responsibility of the whole process. If the instrument does not scan the diagnostic part of the issue, is it now the manufacturer of that instrument who assumes the liability for that misdiagnosis, or is it still the pathologist? And I'm not pretending to know the answer to that question, but I think that it frames the way we think about it very differently.

DR. HEWITT: There's no doubt that it frames the way we think about it and, of course, lawyers look for deep pockets, but going back to the original issue, or question that was discussed, whole slide imaging must offer some advantages because if they don't, there's no reason to adopt it, especially in a situation where if it's inadequate for the question at hand, you have to reflex to glass. We're adding steps to the process. So by definition, they have to bring something to the table in benefit for us to adopt them because the reflex is not to use them, and that's the gold standard. I'm not arguing against the technology. I'm just pointing out the fact that if the

number of examples we have to reflex back to glass, because the technology is inadequate, becomes too high in your work environment, again watch the words, in my work environment, then the technology is not succeeding in its intended purpose. There may be work environments, and we've seen in science the introduction of technology where these instruments work beautifully with their intended work environments in a broad advantage to medicine and improved safety, efficacy, and treatment of patients.

But there are probably times, and we've seen these, where the instruments probably get in the way. We want to make sure that whole slide imaging doesn't become a technology that got in the way but is an enabling technology. So we have to clarify what it can do for us and what it can't, and I think that from my personal experience, having worked with these instruments, without the specifications of what they can do and what they can't do, we can't figure out what they should be doing in a clinical environment because -- I can give you an example that's from my own hospital. I'm not sure they can handle the automated cover slipper that my lab uses. Now, that's not a failing of the instrument. I think it's a failing of the cover slipper, but the point of the matter is, if we've got a mismatch, we've got a problem.

DR. ADCOCK: I think we probably need to move on at this point. So thank you very much to the Panel for that discussion, and at this time, regarding question number 1, which has to do with what U.S. or global

standards, guidelines, industry operating practices are available to guide the characterization of optical microscopy and whole slide imaging, I guess the answer is that the Panel is not aware of current guidelines, and there was robust discussion really concerning what those guidelines should cover, from the pre-analytical to the analytical, and also discussion as to equivalency and what we're really looking at, whether it's equivalency of the optical microscope to the whole slide imaging system or whether the whole slide imaging system is held to a higher standard and is asked to do more.

There's also some significant discussion as to the specifications that the system should take into account, and is it applicable to all aspects of diagnostic pathology?

And then finally, there is some discussion as to, and I'm losing my train of thought, the performance characteristics of the system itself and whether or not we can define specific characteristics for the system.

Dr. Becker, did we answer your question sufficiently?

DR. BECKER: It was an interesting discussion, and I think you have gone as far as you can with it. Thank you.

MS. FAISON: Question 2: Are there objective methods for using the physical phantoms that would be useful in the design, development and validation of WSI in comparison to the OM without the need for human reader interpretations, for example, grids?

DR. ADCOCK: Dr. Gilbertson.

DR. GILBERTSON: Probably. Probably. Most likely. I want to say something. Yeah. I think it's a good thing to look at, yeah.

DR. DAVEY: I was going to say I think we should work -- I mean I think the ideas brought up were interesting. I think the tissue arrays might be interesting for some things. On the other hand, I think having a variety of size and spacing of objects would be very useful to look for completeness of scanning, for example, the problems brought up. You could construct things for that. You could construct I think various phantoms for looking at, you know, the overall imaging characteristics that a pathologist wouldn't have to look at.

So I would think that you could construct something, and it would be useful for promoting some uniformity.

DR. ADCOCK: Dr. Foran.

DR. FORAN: I think if we're not rendering diagnosis in this case, the question is what is the OM? Is it the one that's in John Gilbertson's laboratory, or is it the one that's in his laboratory or the one that's in my laboratory? Each of those would have different sets of configurations in terms of the objectives and so on. So we would have to first standardize the OM before we would start comparing things to it.

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: I think your observation is correct, but we could go back to the example of the CT scanner where we've been able to develop

phantoms that allow them to compare between manufacturers and such, and I think that's beneficial. There's currently an effort to try and develop phantoms that allow them to use the CT spiral or a spiral CT for measurement of small lung tumors and the equivalence between different manufacturers, and I would think that phantoms could be applied in the same fashion.

I think that the diatom phantoms may be very useful in allowing us to understand the effects of disease stack or the depth that the specimen is. I guess one of the manufacturers said 12 microns, 5 microns. I don't know how thick a specimen should be, and the diatoms are something that may really help us understand the resolutions across a 3-dimensional object that the instrument is obtaining because that's going to be dependent on that, numerical aperture, the illumination and numerical aperture of the illumination, and so there's an example of where the phantoms allow us to create a dataset for comparison between platforms, maybe not necessarily against the OM, although at times it may be useful against the OM, but at least it would give us some objective data.

That said, we seem to always ignore our own internal phantom, and I'll put it to you. Most tissue has red blood cells in it, and the red cells are pretty decent human constant, and you can use it as an internal phantom. And I think that that's something that we might -- that I put out there, that it could be used as a reference structure, and interestingly enough, it might allow you to address pre-analytic variables as you move forward as to the

quality of the object that you're looking at because you could define how red it should be. That gives you a color phantom as it were because color is probably the hardest phantom we're facing.

DR. ADCOCK: Dr. Thoms.

DR. MELLO-THOMS: I agree with the use of the phantoms for comparison purposes, but I don't understand how you can use phantoms for validation of new technology as it's encompassed in the question. The only purpose of an image is to be used for something. Technology is there to help the pathologist to diagnose disease. The image per se doesn't mean anything. So I don't understand how the phantom can be used to validate whether it's a good technology or not.

DR. ADCOCK: I'm going to ask a question of Dr. Thoms. Couldn't you use a phantom to determine the operating characteristics?

DR. MELLO-THOMS: Yes, but that's not the validation of the technology. You can use the phantom for several things, and I think using the phantom is a fantastic idea. You can calibrate the system. You can determine what are the characteristics of the system. But validation of the system requires in my opinion the pathologist at the end of it that is going to tell me this system is useful for my diagnosis or this system is not useful for my diagnosis.

DR. HEWITT: No, you're correct. You can only use it really for verification of the system, not validation of the system. I will point out that

you could use the phantom for focus, and in the few experiments I've done on diatom and focus, they're a unique little beast.

DR. ADCOCK: Dr. Birdsong.

DR. BIRDSONG: Well, I agree with most of the preceding comments, but it would seem to me, or at least after you investigate, you might be able to define a set of operating characteristics with phantom and then essentially validate, you know, a set of operating characteristics. So that's not a real clinical validation like I think you're talking about, what you'd have to do with specimens, and that's certainly a part of, you know, developing or defining adequate operating characteristics, and real live specimens always have something that may not be covered in a set of operating characteristics, but, you know, the operating characteristics, they're basically physics, and I think the FDA should at least investigate whether or not they can define, you know, a set of operating parameters sufficient to, you know, constitute, you know, valid operating parameters for a whole instrument.

DR. ADCOCK: Dr. Hewitt or Dr. Zhou. Pardon me.

DR. ZHOU: Well, I think the -- about phantom, you -- gold standards, but if use it through tissues, you might don't have gold standards. I think that the advantage of the use of phantom here, or if you able to use phantom proximate the truth, that's better, but I think use of phantom has a place in validation of the instrument in my opinion.

DR. ADCOCK: Dr. Sinard.

DR. SINARD: I do think that the one thing the FDA could certainly, the one role that would be very useful, if I as a pathologist am looking to buy a new manufacturer's whole slide imager, I would just like to know that I'm not buying a piece of junk, and so it would be nice if there were some sort of minimum, and I guess we're talking again about more operating thresholds rather than validation for diagnosis. But one could imagine creating a series of physical phantoms to address things like what is the contrast threshold necessary for an object to be detected within an ROI sampling software, region of interest sampling? What is the dynamic range of intensity transmission that's detectable by the instrument? What is the response of the instrument to variations in the thickness of the material? I think those could be, you know -- one could image that there could be certain standards that you could say for a piece of equipment to hit the market, it has to meet these thresholds in these particular response range to these phantoms.

DR. MELLO-THOMS: My comments were in relationship to the question posed by the FDA which actually asks about what would be the use for physical phantoms that would help design, develop and validate a whole slide imaging in comparison to optical microscopy. So I'm not the one putting the clinical validation into the question. The FDA put the clinical validation in my perception of the question, and I just don't think that it's the correct way

of assessing these devices.

I agree with Dr. Birdsong that phantoms would be very important to characterizing physical components of the system. So you don't want to buy a piece of junk, but I don't see them as being useful in terms of validation.

I can give you all of the specifications for a -- mammography device. That tells you nothing about how that device is going to perform in your practice. It's going to depend on the readers that are going to use that device.

DR. ADCOCK: Dr. Becker.

DR. BECKER: I'd like to encourage this discussion, please, because it is a core aspect of what I think the Agency is trying to get insight into, that is whether there might be aspects of evaluating the system, short of looking at all 600,000 specimen images that would be necessary to cover the entire breadth of surgical pathology, which can give one confidence about the performance of the instrument as it would be appropriately used with a lesser burden of that clinical validation. So the idea here is to determine whether there is something short of an exhaustive survey of surgical pathology that can be stood in for on the basis of some of these biological or physical measurements.

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: And I think that's exactly the model I'm using is

that the phantoms allow you to develop a verification which allows you to define the specification which would allow you to then certify, I think that's the word language, approve a certain instrument for a certain application, rather than having to use an exhaustive list of diagnostic elements: if an instrument is capable of X, it would be approved for Y based on that specification that it was capable for X. And so I see that as the use of phantoms for verification.

Validation is a separate process. It requires the clinical interaction. As a means of lightening the burden on the validation, basically going back to the assay, once you've demonstrated that the assay works, is reproducible, then you can leverage your interpretative methods on top of your assay as being true and reproducible.

I will say that I think that the introduction of phantoms may be very important in one regard here specifically, and that is the aspect of focus on an image. A couple of the speakers mentioned soft focus. It's unclear to me what constitutes the focal plane of an object when I'm looking at it through a virtual slide. As a pathologist, I go up and down with a Z-axis, and I find what I consider the focal plane. This is an instrument that is doing this in an automated fashion, and as I asked Dr. Descour, in a perfectly flat specimen, that focal plane would be immediately under the cover slip. The tissue is not flat. Tissue preparations, because of stretching everything else, actually have a great deal of undulation and unevenness in the surface, and it's probably

this function that causes the softness in the tissue, and so if you can refine or define what the focal point and the focus features of the image might be, I anticipate that you'll actually be able to better specify what you can resolve and how you can use it.

This is the complexity of a 3-dimensional specimen because if there were flat objects and points, this discussion would really be moot.

DR. ADCOCK: Dr. Sinard.

DR. SINARD: The examples that I gave, and another questions, you know, whole slide imaging in comparison to optical microscopy, I would actually change the word comparison to contrast because the examples I gave for what I see as the role of a phantom is to address those components of the evaluation process that I am giving up by using a digital image rather than the microscope.

So if I have the microscope, I'm responsible for my region of interest selection. If I have a microscope, I can change the intensity, and I can focus up and down. I give those things up when I go to a digital image, and so phantoms could be certainly very helpful in assessing the performance of that device with respect to those components of the evaluation process that I'm going to be turning over to the instrument.

And I think that's where the focus of the evaluation process should be, and so phantoms that specifically see what is the minimal contrast necessary to detect something, as being something that needs to be scanned,

the dynamic range of transmission as well as the responsiveness to variations and thickness, those are things that by using the digital image rather than the microscope, I've sort of lost control over. So those would be the specific areas that I would like to be assured that the instrument is performing appropriately in.

DR. ADCOCK: Dr. Davey.

DR. DAVEY: Yes, and I think along with this, with the development, is just the ongoing maintenance of the instrument, and we have to remember that most of the studies have been done in high quality academic institutions or similar, and if these go out into the community practice settings, there's going to be fewer people that know how to handle them, and they're going to want to know that things are going to work day-to-day.

So I think all of these things that can sort of help calibrate and maintain the instrument because, you know, you can say, well, I want to handle this specimen this way, and if it's a lymphoma, I'm going to want to do it this way, but when you get into a big practice, you're going to be throwing, you know, biopsies, it's going to be processed generally the same way except maybe you have kidney biopsies that are processed different, but you're going to have a relatively standardized way of doing it and, you know, it's not like, you know, academic institutions at that point. So we have to keep that in mind.

DR. ADCOCK: Dr. Zhou.

DR. ZHOU: I wonder whether we should add some specification when you probably answer the question about validations. Validation here is not mean clinical validations, which I talk about those three phases. Maybe this validation here is exploratory stage. We need to specify that. I mean for clinical validation, you can't do it with phantom, but for the exploratory phase or even the challenge phase you can do it. Probably we should have the language here, I think.

DR. ADCOCK: Dr. Gilbertson.

DR. GILBERTSON: I was going to second Dr. Davey and Dr. Sinard's point. In the clinical laboratory, even at these high quality academic medical centers, it's amazing how easy these devices get out of kilter, and you really want to have something that you run every day and you make sure it's running every day, and that would be great. And the real thing that is more important than anything else is that on the part of the analysis, that you've given up, I mean, and I think Dr. Sinard's exactly right about that.

I don't think I want to put a device in my laboratory unless I was really sure that my device, for example, would not miss tissue. And so any device like that would be very useful for the industry because it would make everyone else happy and confident that we're not going to get a lawsuit next week because my device didn't image something.

DR. ADCOCK: Dr. Birdsong.

DR. BIRDSONG: I actually agree with and want to second the points made by Drs. Hewitt and Sinard. That's what I had my hand up earlier for, but also let's not -- since we're talking about, you know, phantoms and standards, let's not write the standards in such a way or try to come up with standards in such a way that we restrict the development of the field like Mr. Bracco said earlier.

So in writing the standards, they should be written in a way that anticipate innovation and improvement so that if someone three years from now comes up with an imager which is adequate for doing hematopathology or cytopathology, we don't have to have this meeting again and discuss all this again because it's really just an extension of parameters that we pretty much already know now because the hardware can't do it but, you know, we shouldn't have to, you know, have reevaluation or the companies that comes up with such a machine shouldn't have to reinvent the wheel so to speak.

DR. ADCOCK: Other comments? All right. At this time, to summarize this discussion, and the question related to are there objective methods for using physical phantoms that would be useful in the design, development and validation of whole slide imaging in comparison to the optical microscope without the need for human reader interpretation, for example, grids, the Panel generally feels that these systems need some measure to verify operating thresholds and characteristics and that these are important not only for the implementation of the instrumentation but for

ongoing maintenance as well.

There is, however, some discussion as to whether this sort of verification is sufficient in that it does not meet the needs for diagnostic validation. So diagnostic validation is a distinct process that perhaps would require different phantoms.

I think the Panel also felt that it was important that the phantoms realized the imperfections or the potential pitfalls of the system and that they tested for such, if possible.

Please, Mr. Bracco.

MR. BRACCO: I don't believe I heard that different phantoms would be used for validation. I believe I heard that no phantoms would be used for validation. Did I not hear that right? Or that it wasn't practical to use phantoms for validation.

DR. ADCOCK: Clinical validation. Dr. Thoms.

DR. MELLO-THOMS: Yes, I agree with Mr. Bracco. I don't believe that different phantoms will be used at all for validation. I think the phantoms are useful for verification of the characteristics of the system.

DR. ADCOCK: Thank you for that point of clarification. Dr. Hewitt.

DR. HEWITT: I can foresee the use of phantoms as an element of the validation. In and of themselves, they would not be sufficient for the validation, but they might be incumbent in the calibration between

investigators in the process of moving towards something for the validation. So very specific, but you could see a role for them, maybe minor role, but they're not --

DR. ZHOU: I think it's validation, the challenge phase I think. Like I mentioned, there is the challenge phase for all the development of the diagnostic tool. So that's -- a phantom could be used for that validation.

DR. ADCOCK: Dr. Becker, is this sufficient for the -- now?

DR. BECKER: It's very helpful. Thank you very much.

MS. FAISON: Okay, Question 3: Are there biological specimens (non-human or human) that could be used to objectively challenge the performance of the human observer in a more controlled way compared to a clinical study? For example, invasion of cancer through basement membrane, fine chromatin details necessary for lineage of hematopoietic cells in tumors and inflammation exudates requiring at least 40X optical objective and Z-dimension focusing, cytoplasmic fine details, etc.

DR. ADCOCK: Dr. Thoms.

DR. MELLO-THOMS: I think this question needs to specify a lot better who the human observer that is being challenged is because if you're talking about an expert pathologist, it doesn't matter, in my opinion, which of these things you're going to use. You are not going to challenge that observer. If you're talking about general pathologists and you give them very complex cases for very specific subdomain, you're challenging them, but is

that a good way to assess the answer to this question? So I think there needs to be a whole lot of better definition about who is being challenged here and what is the purpose.

DR. ADCOCK: Dr. Becker.

DR. BECKER: So we would welcome your observations concerning how that selection of appropriate observers might be obtained.

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: I think there's an overemphasis on challenges in the oncology field without an emphasis on infectious disease and other disorders that we currently encounter in pathology, transplant pathology, hematopathology, such as hepatitis C, hepatitis B, liver biopsies are good examples of challenges that need to be considered within this number of challenges, that one would expect a pathologist to interpret. One could leverage a great deal of knowledge that we've collected through the Board and other mechanisms, the balance of what a pathologist might encounter.

One of the challenges is that a majority of the glass that a pathologist sees is not malignant diagnosis. In fact, it's confirmation of such. These represent the rare birds, and so you almost have to salt the challenge system with boredom and then test them aggressively. If all you're doing is testing them, you don't accomplish the goal because they've got a heightened sense of alertness.

DR. ADCOCK: Dr. Thoms.

DR. MELLO-THOMS: I also think that as part of the way this question is formulated, what is the actual purpose of challenging people at the high magnification? If they pass that test, does that assume that they'll perform as well at low power, at medium power? Because from what I hear from my pathologists, a lot of the diagnosis they do doesn't actually require magnifications of 40X. They do pretty well in like up to 10X or whatever it is that they use. So if they pass this test at high magnification, are they certified in some way, some shape or form, that now they're proficient in the use of this technology? I just don't understand what is the purpose of this test of only one specific resolution.

DR. BECKER: So I would be happy to defer back to the team who actually posed the question if they can clarify further, but I would observe I think that the idea is not to insist that a 40X magnification would be used for all of the testing that would be biological phantom-based, that one would be able to look at the range of magnifications that would be otherwise available for the pathologists to pull in because we know that there are smaller levels of magnification that are essentially constructed from the 40X; but that the team would like to be confident that those features that require in a routine setting a 40X magnification would also be examined and would pass with respect to the biological phantoms that are applied.

DR. MELLO-THOMS: I think that it's a very important point, but would you agree that it's also as important to test them at levels that they see

more commonly, which would be with lower magnifications because if there is an emphasis on fine structures and very high resolution features, perhaps a certain group of pathologists may not do as well on the test but may use the technology very well because that's not what they're going to be doing in their general practice.

DR. BECKER: I'd agree. I think that part of the question is to ask the Panel's help in defining what might be the appropriate scoping of the features that would be looked at with these kinds of biological phantoms. The 40X is an extreme, but I know that in earlier discussions, there have been mention, for example, of pattern recognition that might be approached through biological phantom as well. So that it's across the entire span, but the 40X is just one aspect of it.

DR. ADCOCK: Dr. Sinard.

DR. SINARD: Yeah, I think that this is going to be the toughest part of this whole process to try and crack, you know, because if you're trying to sort of evaluate the performance of the system but, you know, it's such a huge component of that. So basically the analogy we're trying to use is again we can easily fall back on the CP type tests where we have a variety of different specimens that we challenge an analytic system to and see if we can get reliable results at the other end.

Here, you know, the analytic device is still the pathologist, and what we're trying to look at is to what extent, how much variation can be

tolerated in the distortion or artifactual manipulation of the input data that will still give you the same output.

And this process of digitizing the image is only a very small portion of it. There's the, as I mentioned before, the other components in the pre-analytic process, the cutting and the staining. I think that a digital image of a well-stained, well-sectioned micron-based carcinoma will be more accurately diagnosed than a poorly sectioned, poorly stained micron-based carcinoma viewed through a light microscope.

So there are so many other variables prior to the imaging step that come into play, and then an enormous variation in the analytic component, the ability of the pathologist to do that interpretation, but ultimately what it comes back to is, you know, that being the case, you need to take the digital imaging out of the picture at all, why do pathologists -- why does pathology even work if there's that much variability in it? And I think what makes it work is that each pathologist has a pretty reasonable understanding of what they are comfortable with and what they know and what they don't know, and when they need to seek additional testing or additional -- when they need to put adjuncts onto their analytic process, and so there's a tremendous variability in that, but ultimately there still is that pathologist governor to decide when the quality of the data input is sufficient to meet their determination.

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: I'd like to echo Dr. Sinard's comments. I think he puts them in a good context. They argue against Dr. Thoms' arguments, and I think she's not standing for this approach, rather pointing it out, and that is if you go down this route too strictly, you end up creating instances where you are certifying the machine as adequate or inadequate for or you're certifying pathologists as adequate or inadequate to use the technology to make a diagnosis. I don't think that's where we're aiming.

What we're aiming at is learning what the limits of the technology are. I think that's more of what Dr. Sinard is talking about. I would give you the example that I think we're more likely to learn more in those histologic examples in which we grade and where things are not binary. So breast, prostate, transplant kidneys, non-alcoholic steatosis of the liver. Those are just some examples, maybe salivary glands that are used for surgeon's disease. There are tumors, brain tumors. You might want to look at microvessel density in the accuracy, and those are examples where pathologists already are engaged in benchmarking and comparison between pathologists, and I think we'll get a better understanding of the capacities of the technology using that approach than an approach of basically certifying is the pathologist capable or incapable of using the technology.

I do think that, going back to the example of prostate, at some point you have to use the expert pathologist. Well, it's even been done in HER2 testing now. You can use the expert pathologist, the general practice

pathologist, and the resident pathologist. You want to get an idea of is everybody in the spectrum capable of using this technology, or are we running into problems? Now, it's going to be a moving target? Residents may have acquired these skills faster than some of our senior pathologists, but it may be useful for us to know that.

DR. ADCOCK: Dr. Davey.

DR. DAVEY: I just want to basically point out I don't think we can necessarily assume that pathologists always know when they don't know something. I mean I guess I wish that were the case, but that's one of the things we're trying to get at with certification and also going through residency programs and saying whether residents are competent and, you know, all of that. So it would be nice to say that, but there's a lot of variability in that.

I agree with the comments about the types of specimens. We need to look at a variety of specimens, including inflammatory, and I think you could come up with a range of specimens. I think for some of the specific things you mentioned, it's actually not that difficult. I mean for bone marrows, for example, you can take a lot of standard normal bone marrows, and see if the imaging and stain is okay pretty quickly, I feel like, because I look at a lot of those. You don't even need a lot of necessarily malignancies, a few mild -- normals. You could -- I think that's pretty easy to construct later on, but we're sort of excluding it, probably going to exclude it out of those

initially. I think also looking at different kinds of inflammatory cells on H&E, like eosinophils. Can you see the granules and the nuclear -- can you tell an eosinophil from a neutrophil. I mean that's, you know, just an idea of something. Can you pick out organisms, you know, in a stomach biopsy? That's another thing, you know, you can just put a number of those specimens in there which would sort of test the system, and you may decide you can't look. Once you're thinking about an infection, then you really have to resort to glass, and that would be useful to know.

DR. ADCOCK: Dr. Gilbertson and then Dr. Thoms.

DR. GILBERTSON: I kind of agree with this, but with a big caveat, I think. When I teach my students, for example, I say here's a whole slide image and, you know, look at a few things. Look at the height. Can you see through the nuclei? Can you see micronuclei, that kind of stuff? And they can't. They can't see through it. What I never say is it's a bad image. It's a bad slide or an image. The process failed, and it's a big process. I can guarantee that I can give you the same whole slide imager, different cutting, the thickness of a tissue differently, changing nothing else, I can make you fail or not fail those kind of tests. And I can also tell you, that I can either tell you you'll find either eosinophils or not, the biggest after 20X is the staining because if you look at the granules, it's granules usually.

So there's a lot of issues here, and I think what Dr. Sinard said, I think, is correct. The thing that scares me is not, you know, whether or not

they can see a certain thing or not because usually pathologists understand when it's not in focus, it's not in focus. It's very clear, right? It's very clear. If the dynamic range is not enough, it's all black. Pretty clear. But it's those issues, and I think we need to have studies to find what these issues are. The pathologist doesn't know something's wrong and can't tell by looking at the image that something was wrong, and I've talked about a few of these already, but I'm sure that there will be others out there also. We just don't know what they are yet. And those are the ones, I think, we worry about because I am sure in terms of, you know, is this a good device or bad device, that's almost a CAT issue maybe, you know, they evaluate the LAS systems every year, but for FDA the question is safety more than anything else, right, and it's a safety issue, and the big safety issue I still think is when pathologists don't know that they've been misled by the devices, and it's usually not an image issue. It's usually a robotics issue, I think.

DR. ADCOCK: Dr. Thoms.

DR. MELLO-THOMS: I just would like to clarify the comment I made. My point is not that, you know, these tests should be used to make pathologists pass or fail in using this technology. My point is just that I don't understand why we are testing human observers on a technology that we don't even understand ourselves. And so if a pathologist comes in and fails, does it mean that it was the pathologist that wasn't good? Does it mean that the image wasn't good? It didn't contain enough diagnostic information for

the decision to be made. I think until we have more information about how reliable, how reproducible these images are that are generated by whole slide imaging, I don't understand why we are testing pathologists.

So my comment was just in that line, and it was also to clarify that testing with a specific magnification level doesn't necessarily translate to an overall assumption that, you know, if you are good at this, you know, high special frequency element, you'll be good at everything. I don't think that that's the way images are read. So it's just a clarification.

DR. ADCOCK: Dr. Foran.

DR. FORAN: So to make certain that we're not testing the pathologists but we're testing the system, maybe we should start out with just some routine cases where any certified pathologist should be able to make the diagnosis and then move on to the more complicated cases. That would be my recommendation.

DR. MELLO-THOMS: I would agree with that, if we made sure that the way those images were made in these routine cases, that the images had enough diagnostic information that even the most general board certified pathologist could read the image and arrive at the correct diagnosis. So until we have control over how much diagnostic information the image contains, I don't think we should test pathologists.

DR. ADCOCK: Dr. Sinard.

DR. SINARD: I like your point about using sort of routine things,

but I think that the problem with using just routine cases is there is so much excess capacity, if you would, built into the system that, you know, you take most routine cases, and many pathologists don't even need a microscope to make the diagnosis, and so if you really want to sort of test the limits of the system, you've kind of got to get to that other edge, that edge where -- and the question is, how big -- what we're trying to assess here is how big is the zone where the pathologist would feel uncomfortable making a call on a digital image but would be comfortable making a call on the glass slide that was the source of that image. And that's sort of -- we're trying to define how broad that area is.

DR. MELLO-THOMS: Can I comment on that?

DR. ADCOCK: Yes.

DR. MELLO-THOMS: How would you measure that the pathologist is comfortable making that challenging diagnosis on the light microscope? I mean they might not feel comfortable to make that diagnosis. They might want to consult. So how can you even measure that?

DR. SINARD: You can't. I understand that, that that's an issue. I mean you can try to. You can't objectively measure it. You can try to get an assessment of the pathologist's level of confidence in their diagnosis, and understanding that that's a different whole thing than accuracy necessarily. But I think that you can get a sense of it.

Certainly there are cases where -- because it comes up in the

daily practice where I may see something on one section, I'm not comfortable that I feel the section is of sufficient quality for me to call this cancer or for me to call this whatever, and so I'll request a re-cut so that I can, you know, reassess that, and then my confidence level will go up or down.

Now, that's a slightly different situation because I'm getting actually a different image back when I request the re-cut. But I do think that there are models for that.

DR. ADCOCK: Dr. Zhou.

DR. ZHOU: So maybe there's clarifications. We know the system consists of both imaging and the readers. I assume you're going to require company to do that separately or simultaneously, or step by step or -- so let's say first you're going to require the image itself is good. We can use a phantom to do that, the physical characteristic is good. And the next step is to say, well, suppose you have a good image now, how the reader's ability or characteristic affect accuracy of the image systems? Are you think of this way or I think about just do it simultaneously?

DR. BECKER: Well, the intent is not, I'm sure, to evaluate the competence of the reader. Given that you have a certain understanding about the physical ability of the system to deliver an image that meets, let's say, the artificial phantoms, performance characteristics associated with artificial phantoms to a satisfactory degree as far as that can take you, then the biological phantoms would be viewed as an intermediate step to being

able to essentially pick up the features which are not diagnoses but which are elements of a diagnosis. And so the idea here is not to test the readers but to get some understanding about the degree to which the system can bring forward those features that would be appropriate components integrated intellectually later on by a pathologist to help render the final diagnosis. It's not to try to assess the reader. The expectation is that you would find some fashion. In fact, we'd be very happy to understand how we might be able to get this to get readers who would be consistent and competent in their ability to participate in such tests, but the idea is to evaluate the imaging system. The reader is nothing more than the tool that helps to probe that imaging system's performance.

MS. MAGRUDER: I'd like to remind the Panel that please don't respond until Dr. Adcock mentions your name, or if you're involved in a back and forth discussion, say your name first before you speak, because the transcriber is not going to be able to identify the conversation.

DR. ADCOCK: Thank you, Louise. Dr. Kulesza.

DR. KULESZA: I think that Dr. Hewitt already addressed this beautifully, and certainly I don't think that there's any way to compress the pathologist out of this, and I don't think necessarily that -- because what you're really comparing really is the microscope to the image, and you're just using the pathologist as a non-varying variable. So you have to assume at some point that if you give a guy 100 slides with 3 grades of breast cancers,

which presumably would capture the subtleties of imaging -- I mean a guess, an opinion really that we have, but I tend to agree with that very statement, not give them obvious diagnoses but a grading scale, right, that's well established and has certain criteria.

Then if a guy gets 30, 30, 30 on the microscope, but 10, you know, 10 and 90 on the imager, then obviously the imager is wrong. And it's a relative comparison, and it tries to attempt to sort of minimize the role of that pathologist as the variable because it's the same guy, and then it's side by side, or a girl, side by side, the imager versus the scope, and that is the relevant, I think, product of that evaluation.

I don't know, Dr. Becker, that there is anything non-human or human that could objectively otherwise challenge the instrument. I think that the emphasis on the -- is false, and I agree with the rest of the statements that were made here.

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: To kind of bring things together between Dr. Thoms and Dr. Becker, in some ways, we're looking at using this as a means of bringing the specifications of the capacity of the instrument to what the specifications are required for a diagnosis. We're not trying to test the pathologist. We're trying to develop a Rosetta Stone, as it were, of saying what level of knowledge optically do you need to confront this diagnostic problem. It's the only way we can get beyond your exhaustive study

approach is to test this and try and build this. Okay. We know what the machine does mechanically, optically. We know what mechanically, optically that image cognates into appropriate diagnostic histopathology. We have this translator, if we go forward then, a new instrument that met that specification, would functionally be appropriate for that use.

And one would anticipate that there's a stratification of complexity that runs from, to vas deferens, and you don't need much to -- it's a bone marrow, and you can tell the different lineages within the hematocrit, a red cell lineage or something else, and that entire system would work your way all the way down. The pathologist is incumbent on the process because he's the guy that's the gold standard, but the goal in some ways is to try and remove the variability of the faltering gold standard.

DR. ADCOCK: Dr. Davey.

DR. DAVEY: Yeah, I think when you're testing things, it's important to look at the spectrum of observers that would use it. So using all experts or all community practitioners, either one of them is a problem, and I think you have to have, if you use larger numbers, some of the variability is going to be taken care of, and you can, in fact, ask people like on a Likert scale about their level of comfort in using something, and if you have several observers, you're going to take care of some of that. So that's I guess what I wanted to add.

DR. ADCOCK: Dr. Thoms and then Dr. Gilbertson.

DR. MELLO-THOMS: My comment goes again to Dr. Kulesza's point. If you're testing these two technologies side by side, and you are using these readers to determine, okay, are they comparable, you have to allow for a learning curve for the pathologists with the digital slides. The mistake that was made in radiology was to assume that because breast radiologists were used to seeing film, they would just read mammograms the same way once they were presented on a digital screen, and that did not happen. So that was a mistake, and it was a lesson learned that I hope is going to be used here in this context because the image does look different. So all of the perceptual processes and so forth are different.

DR. ADCOCK: If you would like to respond to that, Dr. Kulesza.

DR. KULESZA: Very briefly. I couldn't agree more, and I think that people already made points worth repeating, like you said, about training the person to use the device. An untrained person is going to create noise in the evaluation of the instrument because it's going to be the person. So my assumption is that the person is fully aware of the particulars of the instrument, fully aware of the instrument. Now we're just comparing apples to apples.

DR. ADCOCK: Dr. Thoms.

DR. MELLO-THOMS: Yes. Training the person to use the device is definitely the baseline, but there is also an adaptation of the perceptual and search process that cannot be taught. It has to be learned by the pathologist

himself or herself as they use the technology. So I do agree with you that, you know, somebody has to go and show them, you know, how to push the buttons and so forth, but the learning curve implies mapping, how you learn to read your images in a glass slide that they look this way, this way, this way, to this different domain where the images look fairly different.

DR. ADCOCK: Dr. Kulesza.

DR. KULESZA: So is that then would be similar to a new faculty appearing at the new institution with slightly off stains and having a break in period of a week or 10 days to get the stains, or like Dr. Davey said, the imager pap smears versus the prior. So that could be incorporated into the whole scheme of evaluation.

DR. ADCOCK: Dr. Gilbertson.

DR. GILBERTSON: I think this is a little too simplistic actually. This morning, Michael Becich, he presented some papers, and he saw that with a very primitive device, pathologists could make the diagnosis of breast cancer, all kinds of stuff, just as well as they could under the microscope. Okay. Great. But if you ask the pathologists in those papers, which images were better, they all said under the microscope is much better. I mean, they did. So there's something here that's not -- we're missing something here just saying I could make a diagnosis or not, especially if they're things like breast cancer because I know you can do that without stretching the capabilities of the pathologist's mind or the microscope.

Another way of maybe doing this is doing a continuum, give people 100 images of let's say nevi, a typical melanoma, and see how their confidence is changed as they go back and forth across this continuum. Because I just think that if you give them a bunch of breast cancers or whatever, they'll get that done, I mean unless it's not in focus, where the dynamic range is really out there, and they'll say that, they will get it done. The papers kind of say this where I believe the papers are saying there are certain areas where it gets gray, and the gray area there's confusion, if you look at even prostate cancer where only a couple of glands are available, and if they're a little out of focus, people get confused all the time, but if it's in focus and there's a lot of glands, never a problem.

So I think we're looking at this, is not, you know, can we diagnose, you know, breast cancer with 20X and 1 focal plane, you know. Fine. The question really is how we push that envelope out to where people are competent across an entire continuum of difficult cases. That's not been done before. I know it.

DR. ADCOCK: Dr. Thoms.

DR. MELLO-THOMS: My point is, Dr. Gilbertson, and maybe I'm just not -- from the literature, but it seems to me that a lot of those studies used little bit classified as expert pathologists and, as such, an expert pathologist could probably make a diagnosis in the dark, you know, with the eyes closed, but I don't know for sure that that would be translated to a

community pathologist and so forth. So because the experts could make the diagnosis at 20X with not so good images, I am not quite sure that that would be, you know, exactly the way that the community pathologist would behave.

And my second point is people don't necessarily associate whether they prefer with -- how they perform -- We've done a study where we show radiologists large images and smaller images to go through a slide, a -- stack and they all liked the bigger images the better, but they all performed better in the smaller images. So the way people prefer to look at things doesn't necessarily translate to, you know, what is best for them in a way.

DR. GILBERTSON: I agree with that, first of all. I think these studies mostly were not using expert pathologists, or they were expert pathologists in different fields. I mean -- but I just think that having -- if our job is to figure out whether pathologists can really be challenged by a digital image in a different way than a microscope image, I think we have to kind of push their variabilities a little further out than just saying here's a breast cancer, here's not a breast cancer, because they'll get that. They will get that, and I'm not quite sure that's their job. That's a different issue, but I think that there are other designs than just here are 100 cases, diagnose them.

DR. ADCOCK: One last comment from Dr. Foran, and then we'll wrap up.

DR. FORAN: Yes. I think one thing that was not addressed, are these going to be retrospective studies or prospective studies or a mixture

between the two because you could certainly have somebody reread something that they read five years ago on glass and you have the diagnosis of record, etc., or you could be going ahead prospectively. I don't know what the mixture would be.

DR. HEWITT: Can I --

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: Yeah, we actually did a study that I think that needs to be mentioned because it's germane exactly to that. We did a study looking at cervical dysplasia using virtual microscopy or whole slide imaging, and it was split between two groups. There were 600 images in the cohort, 300 and 300, and in one half of 300, they saw the virtual slides first and then the glass, and in the other half, they seen the glass first and then they saw the virtual microscopy, and they were retrospective slides. And in the group that had seen the real glass first and the virtual microscopy later, there was about a three year gap, and there was drift in the diagnosis, and the other group did not show any change.

So one does have to actually worry themselves about the retrospective nature of these studies because we've demonstrated at least in one instance where there was a drift, and it was probably not from the microscopy but from the change in the grading classifications by the pathologists.

DR. ADCOCK: All right. Dr. Becker.

DR. BECKER: So this is a very rich discussion which I'm very grateful to have heard. At this point, I can say that there is going to be a presentation tomorrow, some information that I think would allow you to rejoin if you choose to, some of these issues, especially concerning reader variability since that's been a very strong part of the discussion here. So that I think you will find there's another opportunity to revisit this to your satisfaction.

DR. ADCOCK: So then as to summarize the discussion, and this relates to the question from the FDA, are there biological specimens that could be used to objectively challenge the performance of the human observer in a more controlled way compared to a clinical study?

Generally the point was made that before any such study is undertaken, we have to make certain that the individual is well trained and familiar with the whole slide imaging technique, and there was also some thought that perhaps this can be accomplished to some level of degree if we use perhaps different types of pathologists, including resident pathologists, general pathologists, and expert pathologists, and if we encompass a spectrum of pathologic tissue types, so looking not only at oncologic specimens but also infectious diseases as well as inflammatory lesions.

And there was also some concern that this might not be sufficient and that we should look at a spectrum of disease with one tissue type, perhaps going from a benign nevus to a malignant melanoma.

Did I capture the discussion accurately? Dr. Becker, is that accurate?

DR. BECKER: Thank you very much. It's been a very helpful discussion.

MS. FAISON: Question 4: What features of image quality (such as resolution, contrast, compression, etc.) are critical requirements and must be achieved for safe and effective use in routine surgical pathology?

DR. ADCOCK: Dr. Birdsong and then Dr. Hewitt.

DR. BIRDSONG: Just a little bit of clarification on the question because the -- you're saying what features are and then in the last part of the question must be achieved, which is our actual specific parameter, not specific parameter, specific measurements, you said specific numbers you would have to hit with your parameters to achieve, you know, safe and effective use, what features, things such as the things that are listed there, but I don't think we can answer the question on this Panel with the information we have so far, what specific numbers we have to hit with the different parameters. We can, I think, give a listing of, you know, what features we need to have numbers on to be able to spec out an instrument so that we're likely to, you know, achieve a safe and effective use.

DR. ADCOCK: Dr. Becker.

DR. BECKER: I think I see the ambiguity in the question and that it, as written, I think, may possibly confuse the features that are to be

evaluated with the performance, with respect to each of those features. The team can certainly correct me if I misinterpret this, but I believe that the thrust of the question is primarily at what are the features that need to be evaluated such that adequate performance, whatever that might be, is assured.

I understand that's correct. Thank you.

DR. ADCOCK: Dr. Birdsong, did you want to speak again?

DR. BIRDSONG: Yeah, just starting with things listed there, with compression, obviously there has to be some compression. I think as a matter of just characterizing the instrument, you have to state both the amount and the compression protocol that you're using since there are several that are available and that might be used. Dynamic range needs to be in the listed, and I feel like I've said that phrase too many times today.

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: This is the one times I'll reflex to radiology. I would say to the FDA, what did you look at with reference to radiology and add color to it? And I think that that will pretty much cover the list effectively, although even now in radiology, with spectroscopy which constitutes in its own fashion color.

That said, I will add an editorial comment, that I don't consider lossless compression an absolute requirement. That doesn't mean lossy compression in and of itself is appropriate, but as one of the figures

demonstrated or in the discussions, luminance may be more important than color features, as we've demonstrated both by the fact that there are colorblind pathologists as well as I'll point out that at least in the space of image analysis, one of the early instruments in the field conducted all of its image analysis in black and white, which points to the feature that luminance is probably a predominantly important feature in these imaging spaces, both for perception as well as analytic tools. So that's the insight that I can add.

DR. ADCOCK: Dr. Thoms.

DR. MELLO-THOMS: Talking about lossy compression, you can have perceptually lossy compression or visually, which would be like visually lossless compression. So you do compress the data, but visually it's indistinguishable from the original image. So that might be the way to go for whole slide imaging, and there are actually techniques to do this. I just came here from a medical image perception conference where they're discussing the use of visually lossy compression on -- slides and they were achieving a range of compression between 5 and 13 to 1, which was not too bad.

DR. ADCOCK: Dr. Foran.

DR. FORAN: I would assume that the tolerable compression will be different across subspecialties. If you're looking at hempath where you have a lot of background, you could probably squeeze out a lot more than you can for let's say a breast or a liver biopsy, and I think that we might have to establish or they might have to establish guidelines across different

subspecialties. It may vary.

DR. ADCOCK: Dr. Gilbertson.

DR. GILBERTSON: I would say that we've been doing images in pathology now for 15 years or more. We've always done lossy compression, usually -- You've had no problems with it really at all within reason, you know, 15 to 1, has always been fine, and also I think that you're much better off having a smaller image that you can transfer quickly, right, than a heavy mess and basically push a lot of data around that you're not going to use. So I think it's kind of a red herring. I don't think it's a big deal.

Having, you know, said that, I think a lot of these issues here depend on what you mean by routine pathology because really it's an important issue. There are cases that you'll never have enough dynamic range, period. It ain't going to happen. Others do quite well on a large percentage of our cases with a lot of stuff we have today. So I think you have to define your practice and then see if your practice, and does whole slide imaging help you in that practice.

DR. ADCOCK: Dr. Sinard.

DR. SINARD: I think there are two words in here trying to achieve safe and effective use, and I think that those are actually kind of different with respect to these features that we're talking about. So the effective use component really speaks much more to overall workflow, and I know that from my prior comments, I've sort of probably given the impression

that I'm, you know, all behind whole slide imaging, and I think it's going to, you know, replace surgical pathology. In fact, I've given talks in actual meetings where I specifically said that I don't think whole slide imaging is going to replace the microscope and the routine surgical pathology practice during my professional career. I could be proven wrong, but I still believe that.

And I think the issue is not because the technology is not there. I think it's more of a workflow issue, and so it was commented on earlier that, you know, if this new technology doesn't perform better than the light microscope and so you have to constantly go back to using the light microscope, then it doesn't work in a workflow setting, and that makes it an ineffective thing, but it doesn't make it unsafe.

And so I think that, you know, people just won't use it if there isn't sufficient resolution, if there isn't sufficient quality such that half the time you have to pull out the microscope or go back to the microscope to make your diagnosis, and people will just stop using the whole slide imaging system. If the workflow is too tedious because the image files haven't been compressed and they take too long to load, and so it takes too much longer to look at the slides, people won't use whole slide imaging. But that doesn't make it unsafe. It just makes it, you know, from the workflow point of view, makes it less effective and that's often the -- step.

But, you know, again what parameters are sort of necessary is

going to vary so much from the type of case that one is dealing with, the differential diagnosis one is trying to make, and also the other variable that I think we need to throw in here is the setting in which the diagnosis is being made.

And, for example, I'll just use if we were talking about a remote frozen section diagnosis, where they need to know now someone's best opinion, and there is no pathologist available locally to look through the light microscope and so it's a matter of either we do this by whole slide imaging or we wait until the next day when the slide can be shipped to the pathologist, then I think that the requirements all of a sudden start dropping because any opinion becomes better than no opinion in those settings. So I think that the type of case, the setting in which the diagnosis is being made and the whole workflow of the environment really speaks to this particular question. The better the resolution, the better the contrast, the faster the performance because of compression, whether lossy or lossless or whatever. Those are all going to affect the workflow and therefore the effectiveness and the ease of adoption of the technology.

Safety is actually a very different issue. Safety is again going to be governed by the question of is the information -- does the digital information contain the information either necessary to make the diagnosis or necessary for the pathologist to realize that they can't make the diagnosis based on that image alone.

DR. ADCOCK: Mr. Bracco and then Dr. Hewitt.

MR. BRACCO: Two things about this question are troublesome to me. The first one is the terminology of safe and effective. That's typically used in the PMA world, and as far as I know, we are still in the 510(k) world for this particular device where substantially equivalent would be used in regards to safe and effective.

The second issue with this question is must be achieved, and my question is what if we have a clinical study that is determined to be statistically valid and is robust enough and it produces results that are equivalent but yet we didn't meet our compression results or compression standard or resolution requirement? Does that mean that the device is not going to be cleared because we didn't meet these? So the word must is a little bit strong, I think, for this particular question because we really at the end of the day, I think, are going to have to rely on the clinical results to make a determination of clearing these devices.

DR. ADCOCK: Dr. Hewitt.

DR. HEWITT: In reference to Mr. Bracco's statement and others and the comment, we want to enable the technology. So if we had an instrument that had limited dynamic range and had a utility, it would have a very limited use. We would like to make sure these instruments have the broadest use they can. This comes back to a comment that Dr. Birdsong made and that is, and I think others have commented or alluded to, and that is that

there's a perception that there's inadequate dynamic range in detection. And so you really can't address any of the other parameters until you've addressed the dynamic range. If you maximize the dynamic range, you're going to alter the contrast functions and the capacities of the compression. Very likely you're going to improve your contrast with a greater dynamic range, although you may not, and you probably will be able to achieve greater compression with a greater dynamic range. You're functionally changing the slope of the bottom to the top when you accomplish this. But it's very likely that it will help you because right now with the limited dynamic ranges, we're running into thick specimens, dark specimens, that we can't get enough light through or cannot perceive enough objects from the light that we have to image it appropriately. If we leverage the system and increase the dynamic range, we've increased the utility of the instrument and hopefully will enable a greater permissive use of the instrument, but unfortunately you have that horse in front, which is the dynamic range.

MR. BRACCO: I'm not trying to downplay any of those, but I'm just, like I said, concerned about the word must. So if the dynamic range doesn't meet a certain parameter but the clinical results show -- and a valid clinical study shows equivalence, does that mean that the device will not get cleared because that particular parameter has not been met, and that's what that word must, as far as I'm concerned, means in that question.

DR. HEWITT: I'm not the FDA, but you're only going to get

probably what you claimed there and would like to enable the technology. As Dr. Sinard said, if the technology isn't robust and useable on a widespread use, people are just not going to adopt it. You want to enable a robust technology because you're talking about a technology that is replacing something, and it's going to have to bring more to the table.

We have an excellent gold standard that's been in use for over 100 years, and so we're talking about a change in practice. It's got to bring something to the table. If it's just good for breasts, well, what am I going to do? Sign out nine cases on microscope and go into the other room to pull up the digital slide? No. That's just not how I'm going to work.

DR. ADCOCK: Dr. Gilbertson and then Dr. Davey.

DR. GILBERTSON: Yeah, I think I have a little disagreement about this. I mean as a pathologist doing your practice, I'll deal with my workflow as I feel appropriate, and if you get a product that will help me, I might buy it. It's really that simple.

In terms of, you know, the must thing, it's kind of an important issue. If I do a -- experiment, and go to my hospital tonight and got -- and take all of their objective lenses except the 4X lens, the next day, there will be a problem, right. The dynamic range wouldn't be too good. The resolution would be pretty bad. We would get some of our cases done, but not many. Then I'll come in and I'll give them a 10X lens, and it will get better. The next time, I'll give them a 20X lens, and it'll get a little bit better. Then I give them

a 40X lens. There would not be a definite change at all. I'm not going to show you. At some point as the field evolves, we'll realize what that dynamic range has to be. I don't think we're going to know it in this room because that experiment has not ever happened for digital slides yet. It kind of has happened for the microscope, 100 years of what we want to buy, we end up buying, you know, up to 40X lenses. You don't buy the 60X lenses usually, right. So I think the must there is an issue, I think, because the experiment has not been done yet. And we can't get the experiment done unless you're actually buying and using the systems.

DR. ADCOCK: Dr. Becker.

DR. BECKER: I'd like to return -- we had a little bit of ambiguity at the beginning of this question concerning the difference between what I think the question is asking, which is, is there an irreducible list of the features that are appropriate to evaluate in terms of the performance characteristics, not what the pass level, for example, concerning those performance characteristics would be. Some of the most recent discussion, I think, is in terms of trying to figure out whether the Panel can recommend how much resolution or how much dynamic range is necessary in order to conclude that the device is performing adequately. That's not the thrust of the question.

The question is in terms of what are the kinds of features that would be appropriate to evaluate in concluding if they meet some

appropriate level of performance, that those features adequately define the instrument's performance to adopt it.

DR. ADCOCK: Dr. Davey.

DR. DAVEY: Yeah, I had just wanted to -- I think most people have been saying that manufacturers need to specify what they're using, not necessarily that we're saying that there's some, you know, limit that you can or can't compress, but the users need to know how things are done, and I just wanted to agree that I think resolution and contrast are more important than the actual color, you know. I think most people agree with that.

One other point I wanted to bring up that hasn't been brought up is that with imaging for cytology now, we're saying on our report, if we're using an imaging system that was brought out at the Bethesda conference. So if we use an imager as part of our diagnosis, we're saying that, and so I just -- that's something that should be -- if we're substituting and we're not using a glass slide at all, is that going to be specified in the report? Just worth considering.

DR. ADCOCK: Dr. Gilbertson.

DR. GILBERTSON: If it's a list of things, there are two that are not there that I think are worth thinking about. One is the ease of navigation, if you will, to get where you want to get on the virtual slide is something that is going to affect how well you make your diagnoses, and we've seen this before. If it's hard to navigate, people won't look at the entire slide. They'll

try to cheat, and they'll miss stuff, and that does happen.

Another thing I think is worth thinking about then is that relate to human factors. Do you get tired of looking at that screen for 18 hours a day? I don't know. And the microscope, we can do this. I don't know how we do it, but we do it. We don't get tired really.

The other thing is that on virtually all of the navigation tools and virtually every -- that you can buy today is a video game. I mean we are essentially playing a video game. You want to go -- if magnification goes up, you click and you point and you click, right. And we're doing this game, and Dr. Zhou a long time ago made a comment like this before where you're taking your eyes off the screen to go and click on something or do something, and that's not true with a microscope. A microscope, our hands do things magically, and our eyes are on the microscope. So I think those two human factors, getting tired, if you get tired, are you going to make a mistake? The navigation issue and the ability or what happens when you continuously take your eyes off the monitor to play this little game, this navigation game? It's worth looking at.

DR. ADCOCK: Dr. Thoms.

DR. MELLO-THOMS: I have a few comments. I think in regards to what Dr. Gilbertson just said, in radiology they've done these studies, and they found that when you're staring at the computer display, your blink rate is reduced, and so you get more dry eyes, and I mean for a long period of time

obviously, and dry eyes leads to fatigue and lead to headaches and all sorts of things. So this is something that is definitely worth looking into.

Also in terms of another point that Dr. Gilbertson made earlier, what he said, that they're using lossy compression at his practice, and that's not really an important issue, but it's important to point out that the FDA requires compression radiology to be lossless. So there is a standard here that has already been put in place but I hope the FDA is not going to follow with digital pathology.

And, finally, I think one point that may not have been brought up about the use of whole slide imaging is that throughout this discussion, we've been talking about pathologists as if they never make mistakes, and we know that mistakes do happen. I think whole slide imaging is a technology that is going to allow us to understand why mistakes are made, and it's going to open door to better training of pathology residents, even like continuous training for current pathologists, and this is going to be done because we can monitor what they are doing, which we cannot do in the glass or the light microscope. So those are my points.

DR. ADCOCK: Dr. Kulesza.

DR. KULESZA: I don't know if Dr. Gilbertson said it, but if he didn't, then I will about the features of the image quality. It's not necessarily quality, but it's the fact that the machine has to scan the entire slide. The slides are standardized, and not to miss tissue outside of the central image, I

think a feature of the image analysis should be checking for parts of tissue that are away from the main specimen or however that machine is scanning at low or high power. Did you mention this because --

DR. GILBERTSON: I did but not recently. So I'll do it again.

DR. KULESZA: go ahead.

DR. GILBERTSON: It's really important. That's the one way you can really hurt somebody. Otherwise, the pathologist has your back in most things, but in that, they might not.

DR. ADCOCK: Yes, Dr. Kulesza.

DR. KULESZA: Anecdotal evidence, and again this was a very curious study which we did with fellows, among the fellows, what were our misses? And actually the highest proportion was not looking at the entire slide. It was not interpretative errors of learning pathologists. It was missing because we didn't look.

DR. ADCOCK: Okay. Dr. Birdsong, and then we'll wrap up.

DR. BIRDSONG: Just a quick, to expand on Dr. Gilbertson's earlier comment with regard to ease of navigation. The FDA might want to think about developing some sort of measure of the ease of finding rare events, which comes down to how well you can navigate, and Dr. Gilbertson may have meant something else when he said it, but when I heard him say that, that's the first thing that came to my mind is how easy is it to find rare events. If you have two systems which are otherwise optically, you know,

similar or electrically similar by whichever parameters are ultimately chosen, but it's easier to find that, you know, small focus of invasion through the basement membrane in A rather than B, you know, or a small micromed in a lymph node, you know, that's a very important parameter.

DR. ADCOCK: So to summarize question number 4 regarding what features are important for image quality, I think some of the features brought up included compression. This may vary by specialty or type of tissue evaluated and compression protocol, the dynamic range, contrast, resolution, ease of navigation, including the ability to find rare events. Also the luminance and the ability of the system to scan the entire slide, and then the other features mentioned by Dr. Gilbertson has to do with fatigue and relating to the operator getting tired and the impact of continually taking one's eyes off the monitor. Dr. Becker, is that adequate?

DR. BECKER: Thank you very much. That's quite fine.

MS. FAISON: Question 5: Discuss the benefits and risks of one z-plane of focus for routine surgical pathology.

DR. ADCOCK: Dr. Sinard.

DR. SINARD: I think the benefits are always there. I mean if you had more than one plane of focus, it would probably ease the adoption because it will be more familiar to the pathologists who want to use the technology. The risks, however, one plane of focus, I think for routine surgical pathology is pretty low. There are very few. I can think of almost none of

diagnoses which are made where we make it on one plane and you don't make it on another plane. I think cytology is a different story where the architecture, 3-dimensional architecture is an important diagnostic element. It's not really the case in surgical pathology. So I think it's a very low risk for routine surgical pathology.

DR. ADCOCK: Dr. Hewitt and then Dr. Gilbertson.

DR. HEWITT: I, in general, agree with Dr. Sinard, but I will comment that the presence of multiple planes would probably significantly decrease the either re-scan rate or the reflexive glass rate for a number of reasons as you work in a routine environment.

DR. ADCOCK: Dr. Gilbertson.

DR. GILBERTSON: I would say if you're going to have one focal plane, which is perfectly good, you have to have humility as a pathologist. You have to be able to say I don't know and I can't see. The issue is not there's something hiding at one focal plane down, but if you're a little bit out of focus, a little bit, that can confuse, and in our studies, it has confused pathologists where they don't know what it is exactly, and they kind of guess because they think it's a study and they have to guess. So you'll see guys that will call cancer, I've seen focal, you know, focal reasons. We weren't able to focus. One guy calls it cancer. One calls it not cancer. One hedges. Beautifully.

So I think the real problem is if you're going to go one focal

plane, which is fine, you've got to realize that there's a limitation there, and there's some things you have to back off. There's cancer everywhere. It's going to be in focus somewhere. So that's not an issue. One small focus of really important tissue, that's a little bit out of focus, that's where you're either going to have multiple focal planes or you're going to have to have humility and say I just don't know. I'll do something else.

DR. ADCOCK: Additional comments?

All right. Then to summarize the question about the risks of one focal plane, I think the committee generally feels that it is limiting, particularly in cytology, and may be an issue when there are areas that are out of focus.

DR. GILBERTSON: With a caveat that that's okay as long as pathologists realize that and work forward.

DR. ADCOCK: All right. Well, at this time, I think what we would like to do since we've gotten through the five questions, this then concludes our discussion of today's questions. I would like to thank the FDA, the guest speakers, and the public speakers for their presentations. I want to thank the Panel members for their attendance and discussions today.

I will remind the Panel to take all your papers and binders with you. Anything you leave behind will be thrown away.

This meeting of the Hematology and Pathology Devices Panel is adjourned for today. We will begin promptly at 8:00 a.m. tomorrow.

(Whereupon, at 6:05 p.m., the meeting was adjourned, to reconvene on Friday, October 23, 2009, at 8:00 a.m.)

C E R T I F I C A T E

This is to certify that the attached proceedings in the matter of:

HEMATOLOGY AND PATHOLOGY DEVICES PANEL

October 22, 2009

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.

_____

TODD BROWN

Official Reporter

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