SESSION ON STANDARDS DEVELOPMENT INITIATIVES TO ENHANCE OVERSIGHT AND ADVANCE INNOVATION OF GENETIC TECHNOLOGIES

Overview of Session

Steven Teutsch, M.D., M.P.H.

DR. TEUTSCH: Now we are going to turn our attention to Standards Development and Initiatives to Enhance Oversight and Advance Innovation of Genetic Technologies. I think, as many of you know who worked so diligently on the Oversight report, control and reference materials play a critical role in assuring the quality and analytic validity of genetic test results. These are the materials we use in performance assessment programs, including proficiency testing.

In the SACGHS Oversight report, we identified a number of significant gaps in the oversight of clinical lab quality and called for stronger CLIA requirements related to proficiency testing and more support for the development of reference materials and methods for assay, analyte, and platform validation, quality control, performance assessment, and standardization.

The National Institute of Standards and Technology, or NIST, and the Centers for Disease Control

and Prevention, CDC, are the federal agencies most involved in addressing these quality control and reference material needs. Currently, reference materials are available for only six of the more than 1,300 clinically available genetic tests. That is pretty amazing, if you ask me.

There are many challenges to the development of these materials, including cost and time involved in producing them.

Given the importance of this area to the oversight system, we thought it would be useful to spend some time delving more deeply into how standards in lab medicine are produced and to explore the challenges and barriers that are impeding innovations in the field and in the translation of biomarker analysis into clinical practice.

We also want to begin to learn about some of the opportunities and initiatives that are under way. We want to explore the impediments to greater private sector involvement and the steps that can be taken to incentivize commercial efforts.

In particular, I would like to thank someone who we hear from regularly, Mike Amos -- who is the ex officio

member from NIST and who has been joining us since I have been on this Committee anyway -- for suggesting the idea of this session to us and, in particular, for helping organize that.

We will start with a presentation from Dr. Willie May, who is the director of NIST Chemical Science and Technology Laboratory. He will provide an overview of NIST's efforts.

Three NIST scientists, Dr. John Butler, Dr. David Bunk, and Dr. Karen Phinney, will present examples of the standards development for genomic, proteomic, and metabolomic tests.

To round out the presentation, Steve Gutman will discuss some of the measurement and standard challenges that are facing FDA, and Dr. Jeff Cossman, chief scientific officer at the Critical Path Institute, will review some of the challenges being faced by clinical labs.

Dr. Amos will discuss future trends in the diagnosis of disease or risk projection, including next-generation diagnostic tests, based on the multiplex determination of complex biomarker signatures rather than single markers of biological activity.

While the focus of today's presentations will be on NIST's efforts, we also want to remain cognizant of CDC's work in this area through its Newborn Screening Program and the Genetic Test Reference Materials Coordination Program, or GeTRM.

We showcased these efforts in our Oversight report. Dr. Lisa Kalman from CDC is joining us today to represent GeTRM. We will have the opportunity to hear from Lisa during the discussion session about the program's current initiatives to develop reference materials for five pharmacogenomic markers and for array-based comparative genomic hybridization, which is a high-resolution analysis of chromosomal imbalances.

Finally, we are also pleased that Penny Keller is here for CMS's CLIA program.

You can find background information on this session at Tab 4 and biosketches in Tab 2. We don't have all of the presentations in your notebooks, but I understand that the remainder will be available to us tomorrow.

Thank you very much, Dr. May, for being here. We look forward to what you have to tell us. Thanks so much.

Initiatives of the National Institute of Standards and Technology (NIST) in Clinical Diagnostics Standards Development

Willie May, Ph.D.

[PowerPoint presentation.]

DR. MAY: We don't have much time, so let's just get at it. What I would like to talk to you about this afternoon is our organization, our basic mission, and some of the new initiatives that we have. Specifically, I will talk about why NIST would be involved in bioscience and health since we are not NIH, we are not CDC, and we are not FDA. I will talk about some of our current activities in the area of bioscience and health.

I will just say now that standards for genetic testing are a very, very small part of the portfolio but one that perhaps you can convince us to increase.

Finally, I will talk about how we are connected to the international measurement standards community.

Our organization was born, if you will, a little bit more than 100 years ago and charged with providing the measurement standards infrastructure to support manufacturing, commerce, and the makers of scientific

apparatus, to work with other government agencies, and to support the academic sector. It is amazing; if you were to look now at the things we do, it is almost like this chart was given to us last year. This still remains the focus of a lot of our activities.

Now, some of the early drivers for some of our activities. We were in the midst of the Industrial Revolution, and people noticed that construction materials were not of uniform quality. Also, there were eight different values for a gallon if you drove from the East Coast to Chicago. Standards were needed for the electrical industry. Scales were not standardized and they were often biased in favor of the seller, as you might imagine.

There were needs from chemical composition, dimensional, and metrology standards to support the railway system. In other words, lots of trains were jumping lots of tracks.

The thing that was most alarming, we being who we are, is we didn't like having to send our instruments abroad to be calibrated. So those things led to the inception of the National Bureau of Standards in 1901.

Since we are not the lead agency for health, the

environment, or food safety and nutrition, and we have this arcane mission of being responsible for the nation's measurement standards, to remain a viable and productive organization we have had to change the focus of our activities continually to focus on major problems of society.

Today our organization has four major components. The NIST laboratories are the remnant of the National Bureau of Standards. We manage the Malcolm Baldridge Quality Award. We have something called the Hollins Manufacturing Extension Partnership and the Technology Innovation Program, which used to be the Advanced Technology Program. Perhaps after the session, if anyone has any questions on any of these extramural programs, I can share those with you.

Our mission is to promote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve quality of life.

If you really were to look closely, this part and that part change. The words change in almost every administration. But these three bullets have not changed

to any substantive effect over the last 100 years.

The NIST laboratories are responsible for maintaining the expertise and facilities for providing this measurement standards infrastructure to support the U.S. That work is carried out by what we call the laboratories, the Chemical Science and Technology Laboratory being one of 10 of these.

As you can see, we are organized pretty much like a university campus. We do what some people might call academic-type research, but that is to support the dissemination of the measurement services products that we disseminate.

Primarily, lots of work goes into the realization of the seven basic units of measurement, things like improving our realization of time. Right now the NIST Atomic Clock is accurate to one second in 30 million years. We are working on clocks now that we think will improve this by three orders of magnitude.

You might think, why would you do this? My watch works fine. Well, things like GPS and a lot of things you don't think about, like interstellar travel and so forth, are very dependent very precise realization of time and

frequency measurements.

The last physical artifact that exists is the kilogram that sits in the basement of the BIPM in Paris. If you have been looking at a lot of the editorials in the popular press lately, you will find that the kilogram is said to be losing weight at about one part in 10^8 per year. We don't really know that that is happening. All we know is that the mass of the kilogram relative to the mass of about 30 other prototypes based on that seems to be changing over time. So the relationship between them is changing, and that is a practical reason for changing.

There are also just pure scientific reasons that are leading the community to try to establish what we call the electronic kilogram. There is an approach to something called the Watt Balance. The new redefinition will be based on Plank's constant, most likely. But to lock that time, we will take this kilogram and then have a device called the Watt Balance. Different countries have different realizations of this to balance electrical force and mechanical force to try to transfer this.

Again, that realization has to agree to about one part in 10^8. Right now, we are about one to two orders or

magnitude off from that. So that has to be completed by 2011 if the kilogram is to be redefined.

But we also serve a much broader community with constantly changing measurement standards needs.

NIST has traditionally focused its research and measurement service activities on the physical science and engineering disciplines. But bioscience and health has now been identified as an area for significant emphasis and growth at NIST.

Why NIST and the biosciences. First of all, as the NIST leadership has looked at our mission, we feel that it is congruent with our mission and indeed our mandate to support U.S. industry and other stakeholders with overcoming measurement standards-related challenges in the biosciences, to provide confidence in results from measurements of complex biosystems, and to enable and facilitate realization of the maximum economic and broad societal benefits of innovation.

Now, Mike Amos and I have this discussion all the time where he says, NIST has to be involved for innovation, and I say, no, we don't, Mike. Not at all. Innovation is going to take place whether NIST exists or not. However,

we maintain that by having this infrastructure to support comparable measurements over space and time we will provide the infrastructure to allow society to gain maximum benefit out of these new innovations.

The other reason that we are doing it is, an emphasis of the administration is a better understanding of complex biological systems. I think this will continue into the next administration. The executive branch, let's say.

Other agencies come to us. This is just one quote. It's from Anna Barker, the deputy director of NCI.

There is an oversight committee that NIST has called the Committee on Advanced Technology. We have heard from two of its members that NIST should also expand its activities to support the biosciences.

Actually, we have been involved in bioscience-related activities for quite some time. Back in the 1920s a collaboration began between NIST and the American Dental Association that led to a lot of the innovations in dentistry that we take for granted now. Things like polymer composite dental fillings and the air turbine drill, found in almost all dental offices, were developed

by a number of employees of the American Dental Association who work at NIST full-time. There are about 30 people. Many people don't know they aren't NIST employees because they work there full-time.

In the 1920s we also started a program in radiation physics which focused initially on X-ray calibration and now includes standards for mammography and radionucleides for radiopharmaceuticals.

We started our program in oncodiagnostics in the 1970s with some support from NIH to provide primary references for electrolytes and metabolites. So, cholesterol, uric acid, glucose, electrolytes, calcium, sodium, and so forth. Then, later, in the 1980s, we began having serum-based standards for those. Around the turn of the century we began to focus on biomarkers for proteins, peptides, and DNA.

This is an example of some of those small molecules, primarily electrolytes and metabolites, that we have had standards for for a number of years. By standards I mean reference measurement procedures and, obviously, certified reference materials or standard reference materials.

Then, about 10 to 15 years ago, we began to focus on more challenging biomarkers. These are some of the things that we have worked on. As you see, two of these might be considered genetic standards, but my colleagues will talk to you about some of the more in-depth details of expansion in this area.

NIST spends a little more than 10 percent of its appropriated funds on bioscience-related activities by our own self-declaration. Now, of this, around $38 million is focused on biosciences. Only about $10 million was appropriated for that. The other money has come as the result of decisions by individual laboratory directors to reprogram funds into this.

Right now, we are in the process of developing a strategic plan not only to support growth of our program in the biosciences but also to do a better job of directing some of the funds that we already have. Right now, to be quite honest, each laboratory has its own program. To get maximum impact out of the resources we have, we are going to try to coordinate this in a much better manner.

I will just go through some of the activities and projects that we have that support health care.

So, what is the typical role of an organization like NIST. We see that all the national metrology institutes around the world have scientifically sound, metrologically-based -- not weather -- measurement science-based competencies and measurement capabilities that are vetted internationally. That underpins the delivery of a number of measurement services, one of which is certified reference materials. Standard reference materials is the NIST brand name for the certified reference materials that we produce.

Now, the Treaty of the Meter was established in 1875. It developed this collegial group of national standards institutes around the world, those that existed. Of course, that was before NIST existed. NIST or NBS, joined that in the early 1900s.

In 1999, though, there was a mutual recognition arrangement that was established that required three things. All national standards institutes like NIST were required to declare and document the measurement capabilities that we use to deliver the services that they provided.

By signing this, you also said that you would

agree to participate in very formal international comparisons so that you had some evidence to support the claims you were making and, further, you would maintain a quality system to underpin your dissemination of the services that you deliver using these techniques that you have claimed have been internationally vetted and compared. This mutual recognition arrangement now has been signed by over 200 national measurement institutes or designated institutes around the world.

This is an example of a comparison for creatinine and serum. This is the European Union laboratory, Korea, the U.K., NIST of course, and the German laboratories. This basically shows how well our capabilities for providing reference measurements for creatinine serum agree with each other

This is a more recent one that was completed this year. This is cortisol in serum and progesterone in serum. Japan, the U.K., China, the U.S., Germany, Korea. Then, progesterone, the same laboratories, except Australia is involved, and Mexico.

In this example certainly, if there was a CRM that was developed by Mexico based on this analysis, there

might be reason to question it, if you will.

The MRA is about documenting measurement capabilities that national metrology institutes maintain and looking at how well those measurement capabilities compare with each other.

Also around 1999, there was this European Union directive that said that the traceability of values of assigned to calibrators or reference materials must be assured through available reference materials of a higher order. The U.S. IVD manufacturers came to NIST and the metrology community and said, we need help with this because without that we won't be able to sell our products in the European Union.

So we convened a meeting at NIST among all the stakeholders. One of the recommendations was the establishment of a global consortium of IVD manufacturers, professional societies, national metrology institutes, and regulatory bodies. This organization became named the Joint Committee on Traceability in Laboratory Medicine. Three principals in this were the International Committee on Weights and Measures, which represents the national metrology institute community; the International Federation

for Clinical Chemistry, which represents the professional community; and the International Laboratory Accreditation Corporation, which represents the accreditation community, if you will.

The product from this is a database of higher order reference measurement procedures, certified reference materials, and laboratories that provide reference measurement services to the clinical chemistry community.

I will just show one of their work products. A work product other than this database is the comparison of standards that are in that database to see how they compare with each other. As it turns out, the standards three years ago for cholesterol came from only two places. There were a number from NIST and a Japanese laboratory, and this just shows how they compared with each other. If one were to select randomly any of the certified reference materials in the database, they agree to within less than 1 percent of each other.

This shows also two reference measurement procedures for cholesterol that are identified in the database, and there are only two. This is how well they agree with each other.

So the world is changing, and we realize that we must change at NIST. Mike Amos is going to talk about this, so I won't say a lot about this except to say that one of the future thrusts for us is to look at tools for what we call visualization of disease signatures and our new initiative for 2010 and beyond. It will have two areas of focus. One is quantitative medical imaging and protein measurement science.

At this point we don't have standards for genetic diseases in there, but after discussing it with you, if the general capabilities that we have won't support that, then there is an opportunity to amend our current plans.

So, thank you for your attention.

[Applause.]

DR. TEUTSCH: Are you happy to entertain questions?

DR. MAY: Sure.

Question-and-Answer Session

DR. ASPINALL: First of all, a very impressive presentation. It was great to give us the history to get to where you are going now. How do you implement new standards? In brief, how does that process work? How do

you get the communication and the time frame to do that?

DR. MAY: Right now we are developing a strategic plan. We are putting together the strategic plan. We have catalogued a number of workshops, conferences, and visits to stakeholder communities. We have captured conversations that we have had when we had official visits from stakeholder communities to NIST to try to develop some sort of coherent plan for NIST.

What we have done in the past is that individual divisions within NIST would conduct their own needs assessment. Lots of the standards that we have now were developed because of input most often from the American Association for Clinical Chemistry. So we would have workshops at AACC meetings often and try to interact with stakeholders and say, what are your top priorities. If you could give us priorities, what would the top five be, for example.

Basically, to answer your question very quickly, we get input from lots of sources. We distill that, try to look at the highest priorities, and then match that with the capabilities that we have. If there is something that is a high priority but we don't have the skill set to

address that problem within the next two or three years, then we tend not to address that because it wouldn't do us any good to have an answer 10 years later when probably the priorities have changed.

DR. ASPINALL: Do you use those same societies to disseminate the information after you have created new standards?

DR. MAY: We disseminate information probably poorly. We have our website. The standards are in our standard reference materials catalogue. Right now, NIST has about 1,400 standard reference materials. About 1,000 of those have values assigned for chemical or biological analytes.

Our old customers know to go through that SRM catalogue to look for what they need. But what we have not done as effectively as we should is provide avenues for new customers and people who don't know about that. That is one of the reasons we are down here today.

DR. TEUTSCH: Julio and then Andrea.

DR. LICINIO: Wonderful presentation. I had a question on the cortisol and progesterone measurements that you had, which was, I think, a fantastic thing to do

because it is true that you have the same sample and you get different measures. It can be very confusing.

One of the things we discussed here before is that one of the issues in the area is that genetic labs sometimes can get disparate results. Would you be willing to do the same type of thing with genetic companies and see what the divergence rate is?

DR. MAY: I guess we could do that. Normally we look to the CAP and other accreditation bodies to do this. This was a comparison among national standards laboratories. These are the laboratories that are supposed to be providing traceability to the companies within their region.

Now, obviously, that is not a perfect thing because right now more than half of the standard reference materials that we sell at NIST are sold internationally, not within the United States. So people are free to get their reference materials from wherever they want.

But this basically is information to the national metrology institute as to how they stack up relative to others. You might ask, how do we know the true answer here? These are not spiked samples. We don't use spiked

samples. We use naturally occurring samples. We have a lot of, let's say, intellectual debates, if you will. We have each of the participants go through their methodology. We shoot holes in it. Then we try to discern from those arguments which laboratories will be used to assign the reference value.

It is not just if you happen to luckily get an answer. We look at the material. For example, LGC's information wasn't used to define this. As it turns out, they were right on. But in their description of their methodology there were some issues. The same thing here. There were only three laboratories that we agreed to consensus had a sound approach.

So everybody develops the approach in their laboratories. This is not using one published method but methods of the highest metrological order as defined by that individual institution. Then we try to get from that to discern what we think the truth is. Then we compare things against that.

DR. FERREIRA-GONZALEZ: Part of my question has already been answered. But, you bring that information back to NIST and assign a value. Before you commercialize

that, do you engage your end users again to see if that value has changed? Do you periodically send surveys out to some of these laboratories to recheck the values?

DR. MAY: It is within our system to do a stability check on all of our reference materials. Some of them might take a year or two years. We might make a measurement now and might make another set of measurements in our laboratories a year or a year and a half later to assure ourselves that the matrix is stable. So it is not until we have addressed all of the issues.

Every certification campaign is different because it depends on what the material is and how stable we think it is. Then we do other measurements to try to assure ourselves that in fact the values are correct and that the material is stable. We do all of that before the customer ever gets the material.

DR. FERREIRA-GONZALEZ: Different analytes for materials will have different times from conception to distribution. What is about a mean time from actual formal distribution of some of these?

DR. MAY: I guess, back when I did useful work in the laboratory I could give you that answer.

[Laughter.]

DR. MAY: It varies so much. For clinical material, I would probably say two years. For a genetic standard, how long would that be, John? A year? I would say a year minimum, probably a maximum of two to three years from the time that we actually began working on the project.

Now, from the time we get input from the stakeholder community, that could be three to four years. Getting the input and deciding that this is going to be our priority, that might take a year's time, because we get lots and lots of input from lots and lots of people. Part of that is deciding internally if this is going to be one of our priorities and making sure that we have the resources to have a successful campaign for development of the reference material.

DR. TEUTSCH: Great. Thank you so much, Dr. May. We are going to take the next three presentations in a row and then get questions after that. Let me turn it over to Dr. Butler, who is going to talk to us about nucleic acid tests.

Nucleic Acid Tests

John Butler, Ph.D.

[PowerPoint presentation.]

DR. BUTLER: Thank you for the opportunity to address the Committee today. You will notice the slides that you have will be different from mine. I will have a few new ones. Some of them will be hidden, so I won't show all of them, in the interest of time.

What I want to show are some of the things we have done in the past and what we are trying to do now with the new Applied Genetics Group that has been formed within the Biochemical Science Division at NIST and within the Chemical Science and Technology Laboratory, and then some of our thoughts for the future.

In terms of the past, most of our experience has come with doing forensic DNA testing, developing reference materials and methods, genotyping assays, and new technologies for improving forensic DNA testing. This is something that has been well noted in the press in terms of the need for good standards and quality measurements.

In terms of the present, two months ago, on October 1st, we formed a new Applied Genetics Group, which is, again, bringing the expertise we have with developing

reference materials for forensic purposes and now applying that to clinical genetics and also agricultural biotechnology efforts, like genetically modified organism detection.

We have some done some work with genetic genealogy and DNA ancestry, trying to help with improving their nomenclature and how testing is compatible within things.

I will finish with just a few thoughts on some planned genetic testing and some of the things we would like to work with. For example, the CDC's GeTRM program. We want to collaborate with them on things.

In terms of our initial efforts and interest in getting into forensic DNA, Congress passed the DNA Identification Act in 1994, which gave the FBI authority to establish a national DNA index system, or national database for DNA testing.

As part of that, there was a DNA advisory board that was formed. One member of that was from NIST. From that came quality assurance standards which now govern how all forensic testing is done in the United States. These standards have also been adopted for testing around the

world as well.

Standard 9.5 within the section on analytical testing says specifically that the laboratory shall check its DNA procedures whenever a change is made against an appropriate and available NIST standard reference material or a standard traceable to a NIST standard. This is what has driven most of our efforts in forensic DNA testing, trying to provide information that can help with the underpinnings of quality measurements for forensic laboratories.

This is a new slide here that I just added showing that at the highest level, the community level, there are quality assurance standards to make sure that there is also, of course, inter-laboratory studies to make sure that everybody can talk to each other in terms of their data.

Within the laboratory, there is the American Society of Crime Lab Directors Laboratory Accreditation Board. They have accreditation of laboratories. Audits are performed, usually annually, of laboratories to make sure that they are compliant with the specifications there.

Each individual forensic DNA analyst must perform

two proficiency tests per year on any type of testing that they are doing, plus they are required to have continuing education to keep up with new technologies.

The next level is the instrument or the method level, where we have validation of analytical procedures. This is where the NIST reference materials come in. You have a traceable reference material to make sure that your instrument or your method is working properly.

Next is at the protocol level, where you have standard operating procedures to make sure that the instruments are used consistently from analyst to analyst and so on. Each data set has its own standard materials that are run, positive and negative controls, and so on. Allelic ladders are a mixture of DNA samples to show all the possible alleles that would be seen.

Individual samples have internal size standards that are run with them. Then we have interpretation of results that are confirmed by a second analyst. Finally, of course, when you go to court, you have defense attorneys and defense experts that can examine your data as part of discovery requests. That provides another check and balance on how forensic DNA results are done.

So, all the way from the community level to what is presented in court there are checks and balances with things. The reference materials that NIST provides are only a small piece of the validation of the analytical performance of something.

Over the years, there have been a number of different technologies that have been used. For each of these different technologies we try to have a NIST reference material available to help with this. The first is, of course, the restriction fragment link polymorphism, developed in the late '80s. That was the initial DNA fingerprinting or DNA typing that was developed.

Then there became polymerase chain reaction-based tests. The next series of reference materials was SRM 2391, which has been available since the mid 1990s. Then we have had ones for DNA sequencing and mitochondrial DNA and, most recently, for Y chromosome testing.

The technology in some cases is no longer used and therefore reference materials get phased out. Then there are growth areas in terms of new markers and new information that can be added to the same samples and certified on the same samples.

This is just to illustrate what we do on the genetic tests. On the top right, you see a picture of the DNA samples themselves. There are 12 different samples that are provided for this particular test. Then there is a certificate of analysis that provides genetic data for each of those samples.

In this case, they were characterized for 22 autosomal, short-10 and repeat markers that are used in forensic testing around the world. We have just recently added 26 new STR markers. It is basically a value added to the same reference material. So the DNAs haven't changed. We have just added more certified information to them.

We have also tried to encourage the slowing down of the consumption of these because they are expensive to make and certify. We tried to help laboratories make traceable materials instead of just using straight off the shelf the reference materials themselves.

These are the basic steps in forensic DNA testing. You collect the sample, you extract the DNA and quantify how much DNA is present, perform a multiplex PCR application. Then you look at the short tandem repeat markers and interpret those results, and then put those

results in a database where they would be checked against the frequencies of alleles to determine how common that particular profile is. That is what would be presented in court if they match.

So the reference materials only focus on the actual typing results that are produced. There are many other aspects of the process that could have reference materials, but right now we are just focusing on the separation of the DNA itself.

We are looking at short tandem repeats. That is what is used in forensics where we have primers that target a repeat region. The number of repeats is then converted. The overall size of the PCR product is measured and then the number of repeats is what is actually considered in the final analysis and what is reported. In this case, 11 GATA repeats is what is recorded in the database for that DNA profile.

That measurement is made against an allelic ladder, which is a mixture of alleles. You can see in this case, just showing two samples, one that is a 16/17 and one that is a 15/16. Both those samples are compared against an allelic ladder that a commercial manufacturer produces.

They check that allelic ladder against the NIST reference material.

There are different sites that are used throughout the human genome for forensic testing. In 1997 the FBI defined 13 core loci. There is also a sex-typing marker that is used called amylogenin that is present on X and Y. Then there is some overlap with Europe. So our reference materials are also used in Europe, though they use slightly different genetic markers for their testing there.

Now, within the U.S. we have over 6.5 million profiles on the database. A laboratory cannot put their results on the database unless they have run a NIST SRM to make sure that their results are accurate and so on.

Again, a little bit more on the STRs. We are measuring the base pair size, converting that back to a repeat number, and that is what is being stored.

This is also used for paternity testing. Our reference materials are used to help with making sure that paternity testing is done properly. The American Association of Blood Banks, AABB, is who oversees how paternity testing is done.

This is what a full DNA profile looks like, just to illustrate the process. An internal size standard is run with every sample. Then we have the individual samples compared to an allelic ladder to actually get the genotypes for each individual site. The measurement is performed by the allele size.

Another thing that is important to point out, of course, is that different genetic tests may use different PCR primers and therefore, because of binding site mutations, may produce different results because of allele dropout or null alleles. This is just to illustrate one example with a NIST SRM 2391b.

The Genomic DNA 8 actually has a dropout at this marker on chromosome 16 with a new kit that just came out from Applied Biosystems. You lose Allele 11. This becomes important as laboratories are trying to verify if their procedure is working properly. So we go through and do a lot of work to calibrate and sequence the regions and define why a particular new assay or kit doesn't work properly.

We are funded primarily by the National Institute of Justice to do this work, as well as internal NIST funds.

We have reference materials, as I mentioned. We have standard information. We have conducted a lot of interlaboratory studies. On the technology side, we are constantly developing new assays and new software. We have training materials. You can go on our website, which is the STRBASE website, and download PowerPoints and other workshop information to help people learn more about this.

Just to get to where we are now, you will hear about some work going on in the Analytical Chemistry Division in just a moment. We are within the Biochemical Science Division. It is all underneath CSTL. We just, as recently as two months ago, formed an Applied Genetics Group, which is one of six groups doing work with genetic testing. These are the people that are involved there. Marcia Holden and Ross Haynes are new additions to our group, the former forensic group. We are really expanding in this area.

Our mission is to advance technology and traceability then with quality genetic measurements, continuing to help the forensic testing community but also clinical genetics, the ag bio tech, and then also DNA biometrics. There is a tremendous interest in this area

and speeding up the process of DNA testing and making sure that is done accurately by the intelligence community, and so on.

This is some of our group expertise and funding sources. We have primarily, again, expertise in reference material characterization, construction of new assays, a lot of work with sequencing, SNPs, STRs, and so on. Our primary funding is coming from NIJ, but we are also getting internal funding from NIST. We plan to strengthen our portfolio in the clinical genetics area.

DR. TEUTSCH: Dr. Butler, I hate to interrupt you, but we will need to wrap this up so we give everybody a chance.

DR. BUTLER: That's fine. These are our reference materials that are available right now. There are some slides from Mark Salit here on some of the RNA work that he has been doing.

We have been trying to help with nomenclature to help the genetic genealogy community to make sure that they are getting consistent results across laboratories.

This is one of the new ones. We are working on Huntington's disease, trying to have alleles that

appropriately define each of the characteristics you would expect to see with Huntington's disease.

We have to decide, and we welcome input, in terms of what types of materials should we certify. We can certify for a sequence, a specific genotype, and of course, the quantity of DNA that is present.

We want to continue making information available to the public, as we have with our forensic stuff, and make that available for clinical diagnostics as well. Feel free to contact me if you have questions, and thanks again for your attention.

[Applause.]

DR. TEUTSCH: Thank you. I hate to rush you through all of that, but I want to give everybody else a fair chance.

Let's move on to Dr. Bunk, who is going to talk to us about proteomic tests. Welcome.

Proteomic Tests

David Bunk, Ph.D.

[PowerPoint presentation.]

DR. BUNK: Thank you very much. Thanks for the invitation to come speak to you this afternoon. Now for

something slightly different, some protein work that we are doing at NIST. This is a new effort in terms of helping to standardize and improve the measurement quality of proteomic clinical research.

Proteomics has not yet moved its way into the clinical diagnostic lab. I'm sure it will be entering soon enough. Right now proteomics is mostly used for medical research and medical diagnostic research. But the important thing here is that the measurements still need to be standardized. There still need to be high-quality measurements in order to make sure that the medical research is moving forward in the right directions and not leading down the wrong paths.

Just a quick definition in case we are not familiar with what proteomics is. Proteomics is the identification and quantification of all proteins of whatever sample you are talking about, whether it is the human proteome or specific tissue proteomes.

The interesting thing about proteomics, where it differs from genomics or metabolomics, is that very little research in proteomics actually measures intact proteins. You can divide proteomics into two distinct approaches:

the top-down proteomics, where intact proteins are measured, but the vast majority of proteomic research is done using an approach called bottom-up proteomics, in which proteins are degraded down into peptides and peptides are measured. Then we are relating that information back to try to figure out what is going on at the protein level.

That is important when we talk about how we standardize the measurement techniques because we need to know what is going on. If things are not being done at the protein level, then we don't necessarily need reference materials at the protein level. We can actually do a lot of work by having peptide-based reference materials.

Clinical proteomics is a subcommunity of all proteomics. Really, from my understanding, the goal of clinical proteomics is to discover new diagnostic biomarkers. It is both looking at the change in the structure of the concentration and interactions with different proteins in order to improve clinical diagnostics.

If we look at the clinical biomarker pipeline, the first phase of biomarker work is the discovery phase, where we identify candidate biomarkers. That moves into

the verification of these candidate protein biomarkers and finally into clinical validation. Currently, proteomics is being used in the discovery phase and the verification phase. The clinical validation is large-scale, large cohort studies in which most of the work is done using traditional techniques like amino assays.

But there is some belief that proteomic measurement technology will be used in clinical validation in the near future, and some of these technologies are being developed in order to do that. But currently, proteomics is focused on the discovery phase and the candidate verification.

The distinction here is, in the discovery phase we are only talking about a small number of samples, maybe one healthy and one disease state samples. As we move into verification, we want to try to reduce the number of candidate biomarkers down to a manageable number, and so we use a larger amount of clinical samples. Of course, with clinical validation, we are talking about thousands of patients in order to make sure that we have a true biomarker that has either diagnostic or prognostic utility.

Proteomics is still in its infancy, to a certain

degree. There are a lot of problems in proteomic measurements. That is one of the reasons why NIST is involved. We want to bring a higher level measurement quality to proteomics.

Basically, I think one of the fundamental problems in proteomics now is that there are no quality metrics. There are no performance criteria. At least, there have not been in the last few years. There have been a number of studies published. The Human Proteomics Organization has published a number of studies where they are looking at interlaboratory comparisons of proteomic investigations. Unfortunately, many of the results are not very positive. There has been very little comparability in proteomics investigation from laboratory to laboratory. Obviously, if you want to develop technologies for doing clinical diagnostics, the field of proteomics had to be improved in order to get more reliability and more comparability of the measurements.

The other issue is, it is very difficult to assess truth in proteomics. No one knows what the human proteome is. It is very difficult right now to assess agreements if you don't have standards. That is one of the

reasons why we are here at NIST.

Unfortunately, all of this has led to the potential of diminishing opportunities for future research funding. On that note, a few years back we partnered with the National Cancer Institute on one of their initiatives and really discussed this.

One of the fundamental approaches we take in developing reference materials and reference measurement procedures for clinical diagnostics is partnering. We at NIST are not clinical chemists. I am not a clinical chemist. What we do know at NIST is the basic fundamentals of measurements.

So what we have to do is partner with professional organizations like the AACC, the IFCC, and the National Cancer Institute in this case, to bring their expertise into our efforts in standardization. We apply our measurement skills, our knowledge of the fundamentals of measurement, and we bring in their application knowledge to solve the problems that are relevant to them.

The National Cancer Institute, about three years ago, developed a program to assess proteomic technologies because, basically, their advisors were telling them that

they are not going to be funding much future research for proteomics because there was no payoff. So NCI decided they needed to initiate a program to evaluate the technologies.

It is a very interesting program because it is not about biomarker discovery. It is about validating the technology used in clinical proteomics.

The role that NIST plays in this program is that we are advising them in some of their interlaboratory study designs and developing the materials that are being used in interlaboratory studies. We are working with them to really help assess the technology ourselves. In the meantime, we are learning a lot about proteomics. So we are gaining the knowledge from the community by working with these partners, and that is an important aspect.

Through this initiative we are working on interlaboratory studies but we are also developing the information we need to develop our own reference material program to support proteomics.

Let me go back to the biomarker pipeline once again to draw some distinctions here. Biomarker discovery is mostly a qualitative or relative quantitative

measurement. This work is mostly done these days in tissues, so we are looking at the sources of disease, like cancer would be in tumors.

The verification stage is doing more of an absolute quantification of signature peptides from whatever the candidate biomarkers are. That is being done in mostly plasma because this is leading toward a more diagnostic platform. The instruments being used are much more qualitative.

Realizing that proteomics is playing a role in both of these fields, discovery and verification, NIST is developing reference materials to support both efforts because if you are not supporting the entire pipeline you are still going to run into problems. We need to have reference materials and standard operating procedures and validation tools for the entire pipeline.

Let me just mention some terminology we use in terms of reference materials, which is horizontal versus vertical standards, or vertical reference materials.

When we are talking about a very complicated measurement technique or measurement pipeline like in proteomics, where there is sample collection, sample

processing, instrumental analysis, and data analysis, there are a lot of places where problems can come in. We approach that we take at NIST is to develop horizontal standards, which are standards which support measurement quality in individual steps along the way.

The other thing we also develop is vertical standards, which are very much application-specific standards.

A horizontal standard might be a standard that can be used to validate your data analysis, whereas a vertical standard would be a more complex, application-specific standard like cholesterol in serum, where it is geared towards a much more specific measurement problem. The standard is carried through the entire measurement process.

In proteomics, that is the approach we are taking. We are developing horizontal standards and vertical standards in order to support the measurements.

In most cases, for a new measurement area it would be impossible to develop just vertical standards. The applications where proteomics is being used are very significant, so we would have to develop vertical standards

for every specific application.

In clinical diagnostics, we have reference materials for cholesterol measurements, glucose measurements, creatinine measurements, and so on and so forth. That approach for proteomics just wouldn't work because there are too many areas in which it is used. So a horizontal standard is a way that we apply our resources to improve the measurement as best we can.

Currently, we have two reference materials in production. The horizontal standard is a mixture of synthetic peptides, so it is not application-specific. It is designed to improve quality in mass spectrometry instrumentation. So all fields of proteomics that involve mass spectrometry could benefit from this reference material since this is a common point in their pipeline, making that a horizontal standard.

The other reference material we are currently developing is a yeast proteome reference material. This is a vertical standard, so this is designed for proteomic investigators to take a complex protein mixture through their entire proteomic pipeline and validate the procedures that are being used here.

We also have plans to develop more complex proteomics reference materials that are plasma-based for quantitative measurements.

In addition to those two new reference materials and the additional one that I mentioned of complex-matrix horizontal standards and vertical standards, we are also looking at developing higher-order measurement tools for assessing performance of affinity reagents in proteomic arrays, multiplex arrays, as well as developing and validating novel affinity capture reagents. So we are looking at both improving technologies, developing standard operating procedures for people doing proteomics, as well as delivering services through reference materials, which people can use to validate their technologies and their techniques in proteomics.

We hope that by having all these different areas we can support the measurements that are going on in the clinical community and improve the outcome of clinical proteomic research.

Thank you.

DR. TEUTSCH: Thank you, Dr. Bunk.

[Applause.]

DR. TEUTSCH: Now, metabolomics. Dr. Phinney, welcome.

Metabolomic Tests

Karen Phinney, Ph.D.

[PowerPoint presentation.]

DR. PHINNEY: Thank you. I'm very happy to be here today. I appreciate the invitation. For those of you who are unfamiliar with metabolomics, this is something that has been going on in clinical chemistry for a long time. We have been measuring small molecules like glucose and cholesterol as part of diagnosing disease. To a great extent, this is just a fancy name for something that has been going on for a long time.

Metabolomics really represents the endpoint of genomics and proteomics. It is what you really get when you look at a sample of serum, plasma, or urine. Those samples reflect the exact processes going on at that period of time.

There are some advantages to looking at the metabolome. It does represent an exact picture of the situation in the body at that point in time, and it is affected by things like diet, stress, exercise, disease,

health, you name it. So instead of looking at the genome, where you look at what might happen, you actually look at the phenotype or what really did happen. To a great extent, this could be the ultimate in really doing disease diagnosis.

There are some other things to know about the metabolome. It is simpler than looking at either the genome or proteome. Even though in the metabolome you are still talking about thousands of potential metabolites, that is still a far simpler situation than thinking of hundreds of thousands of different proteins or even tens of thousands of different genes.

So, what is the goal of metabolomics. Why are we throwing around this fancy terminology. As I mentioned, we have been using metabolites as diagnostic markers for a long time, but we have tended to do them one at a time. We might look at glucose to diagnose diabetes and we look at cholesterol to look at risk of heart disease. But we haven't put all those pieces together. So what is unique about metabolomics is that it involves looking at panels or signatures of different analytes and their levels under different circumstances in the case of health or disease.

Ideally, you can use those patterns or those signatures to try and segment people into different groups and, ideally, use that as a way of doing disease diagnosis.

If you look at the picture that is there on the left, that is an NMR pattern or NMR analysis of a particular sample. You can see there are lots of different peaks there. You can see, looking at the different color of spectra, that there are some differences in how those appear.

The goal of metabolomics is to try to look at those different patterns and to be able to say something about different levels of particular metabolites representing some signature. So, does it represent a healthy person or a diseased person.

Ideally, we would like to get to the situation that you see on the right, where you can put people in different boxes and say in this particular population we see this signature or these different metabolites at these particular levels and in a healthy person we see a different pattern. If you can do that with some reliability, you could use that as a diagnostic tool.

Now, one of the reasons to do this is also to try

and identify places where we could intervene in a disease state. If we know that in a particular disease certain metabolites were elevated or decreased, we could then try to intervene in that particular metabolic pathway through pharmaceuticals or some other therapy. So metabolomics does represent one potential mechanism to identify new therapies, and there is certainly a lot of activity in this area in the pharmaceutical industry.

The drug industry is also interested in looking at this as a mechanism to identify toxicity. If you can identify particular markers that indicate liver toxicity, for example, and you can measure those in a multiplexed way, you might be able to predict ahead of time whether a particular pharmaceutical is going to have adverse effects.

That would certainly be very valuable. We know these days we hear a lot in the news about things that make it onto the market only to be withdrawn later. Certainly, that is why the pharmaceutical industry has such an interest in this area.

Finally, as you saw in one of the first slides there, all these things are related. The metabolome can be traced all the way back to the genome. If you look at

patterns of metabolites, you might be able to say something about gene function that assumes something about the metabolome, the proteome, and the genome all at the same time. That is quite a lot of information to try to capture, but under ideal circumstances you might be able to do that.

So, what are some of the issues. Where does standardization come in. If you think about trying to measure thousands of metabolites simultaneously, you are talking about very large and complex data sets. As David mentioned, there are always issues in terms of instrumentation, sample collection, and sample handling. So, how can you get to a point where you can say with some certainty that the pattern of metabolites that you see is really representative of a particular condition.

There are a number of these issues: sampling, instrument variations, platform variations, and software, just in dealing with these very large data sets.

Once you get your data, how do you pick out which things actually mean something. There are thousands of metabolites but maybe only three are relevant to the particular condition that you are studying. This comes

down to software and it comes down to making assumptions about the data that you have. Clearly, in those situations there is room for error and there is room for differences in interpretation.

Finally, before we can get to a clinical diagnostic setting, we need to actually validate that the patterns of metabolites we think are useful in diagnosis really are. Certainly, that comes back to looking at large populations of people and making sure that you really can say with some certainty that you are making an accurate diagnosis based on this metabolite signature.

About two years ago, I guess, NIH came to us. They have been funding a number of investigators for metabolomics technology development. But along with that effort they realized the importance of some standardization and some common way for people to evaluate the technology that they were developing, some common mechanism for them to use. So they approached NIST about developing reference material for metabolomics.

We have been involved in that effort over about the past two years, and this material will be introduced I think probably early in 2009. So we are coming close to at

least the end of the first stage of this process.

This reference material is actually a plasma pool. The reason that we did that is we didn't want to represent any particular part of a population. We wanted this to be indicative of a mix of male and female, different age groups, and healthy individuals, and we wanted it to also have some of the ethnic characteristics of the U.S. population. So the samples that were pooled to prepared this material came from African Americans, Asians, Caucasians, and, again, both male and female individuals.

One of the reasons that we did that was that when we have to prepare this material again in, say, 10 years, we wanted to be able to prepare it in a very similar way. That is why we set these criteria in designing the material.

We have a lot of experience in measuring individual metabolites. As Dr. May mentioned, we have a number of different reference materials for individual metabolites in serum, the traditional analytes like cholesterol, glucose, and creatinine. We have measured those same analytes in this particular reference material, so we will have certified values for probably 40 different

metabolites, everything from fatty acids to glucose, to hormones.

But we also realized that people want something more than that. They would like to know what other metabolites are present. So the effort that we are focusing on right now is more of a qualitative effort to see what techniques do we have available, either at NIST or through collaborators, where we can identify additional metabolites and also provide that information.

Clearly, there is the potential to use this material in a variety of different ways. Depending upon your particular study, if you are looking at glucose metabolism or if you are looking at kidney disease, your interests may be different. So in order to make this material relevant to as many different people as possible, we are trying to provide as much information as we can.

Now, clearly, this is a starting point in terms of providing standards for this particular area. It is an evolving field, and we certainly recognize that. We do see the potential for additional reference materials and different standards here, and also tools in the area of bioinformatics. One of the big questions here is how do

you handle these large data sets. How do you insure their reliability. How do you compare data from different instrument platforms or different laboratories. I think these are all questions that will be coming up as this field moves forward. It is still very early on.

We also realize that there may be a need for reference materials to focus on more specific populations. It might be a group of individuals with heart disease or it might be male versus female. The list could go on and on. Certainly, we look to the field to help us in prioritizing those efforts.

There are some fledgling standardization efforts in this field, particularly in the area of data reporting. So we are also working with those organizations to offer our insight into metrology and to learn from them in the areas where NIST can contribute in terms of standardization.

With that, I will close. I know we are going to have time for some discussion here at the end. I appreciate your time.

[Applause.]

DR. TEUTSCH: Great. Thank you very much. I

think what we will do is continue on. Then we can take questions at the very end.

Steve, let me welcome you. Again, thank you for all your service to FDA and to the Committee in so many ways, and not only this Committee but our predecessor. Thanks so much. You will be talking to us a bit about the regulatory agency perspective.

Regulatory Agency Perspective

Steve Gutman, M.D., M.B.A.

[PowerPoint presentation.]

DR. GUTMAN: I can't think of a better swan song than to stumble across this topic, so I thank you.

FDA has a longstanding interest in standards. In fact, the original regulations in FDA for our primetime submission, the 510(k), which is what we use for me-too devices, call for the use of standards in equivalency decisions.

In the early '80s FDA initiated development of standardized, traceable methods and expected thresholds for both glucose and hemoglobin, took them to the public, and I guess they weren't ready for primetime yet because we couldn't make the sale.

So what we resorted to -- and in fact the regs were subsequently changed to accommodate for the nascent life of standards in the '80s -- is we changed the regs to call for special controls.

Our program is largely based on two operative terms for me-too devices: showing that they are substantially equivalent to a predicate and, for novel, high-risk devices, showing that they are de novo, safe, and effective. Neither of these regulatory submissions actually calls for or requires identification of either standards, traceability, or performance against standards. I would argue that that is a weakness in our regulatory toolbox.

That has, of course, not been a deterrent to our renegade workgroup. We continue to rail for standards. FDA was a founding member of the CLSI. We are an active member of the ISO Technical Committee 212, an active member of the IBD Subgroup of the Global Harmonization Task Force, and an early proponent of the CDC's Standardization Program. So the lack of standards does not demonstrate a lack of enthusiasm on the part of our workgroup.

In fact, if you bother to look at our webpage,

you can see that when we write guidance we frequently reference standards. When we develop special controls, we frequently reference standards. In fact, if you look at our decision summaries, the more "with it" companies will in fact reference standards.

We also have an interest in the material standards that NIST is developing. We always attempt to identify usable standards, whether they are NIST, whether they are CDC, whether they are WHO, or whether they come from other legitimate sources. We have experience with the use of material standards in both pre- and post-market programs.

In terms of the formal process, there is a formal recognition process, at least for methods standards. About two dozen members of my office participate actively. We have recognized a number of CLSI standards and a smaller number of ISO standards. They are all, again, found on our webpage.

There is a formal process that these standards, once recognized, can be used in the context of pre-market review. There is a particular entity called the abbreviated 510(k), where companies can actually conform to

standards. That increases the certainty and decreases the negotiation between FDA and the sponsor submitting that particular standard.

In point of fact, there is usually partial rather than complete conformance. The CLSI standards are an interesting hybrid, some more geared towards laboratory practice and manufacturing practice. It would be fair to say the abbreviated 510(k) is not a perfect program.

I would also point out that informal use of standards is very frequent. Often pedigreed materials, sometimes from CDC, sometimes from WHO, sometimes from other sources, may actually carry a floundering company over the threshold in terms of pre-market review. While our pre-market review has, I think, weak regulatory tools, the quality system regs that are part of our post-market compliance program do in fact have very beguiling portions of the regs that might speak to. if FDA were aggressive in the pursuit of those regs, the use of standards. So there are interesting tools to look at in the future if there was a call for better standardization products.

There certainly are incentives to do this. The IVD directive in Europe very explicitly calls for the use

of standards. Our transparent posting of decision summaries provides a reward for use of standard materials or methods because it becomes a matter of public information. I would argue the STAR*D initiative and other efforts to provide clinical standardization will only be as good as the ability to have an underpinning of analytical standardization as well.

That being said, there is a long journey ahead. The truth is the status quo for routine assays -- PSA, troponin, d-dimer are three of my favorites -- is absolute noncongruence. If you look at proficiency testing surveys, you will be astounded by the laboratory and company differences. You can get a heart attack simply moving from one ER to another.

The status quo for new assays is worse because there is no proficiency testing. There is no QC material. It is gratifying to see that NIST is starting to move forward, but there is a mountain of new assays, some of them protected by IP, that might make it very difficult to create cross-lab standards.

This has all been further complicated by the fact that in the year 2009 we actually get it in terms of the

complexity of sample procurement and the whimsy of pre-analytical systems in terms of impacting the results any particular system might generate.

At the end of the rainbow, there is a pot of gold. I think Mike may talk about this in more detail. There is a shift towards evidence-based medicine, even laboratory medicine.

Thank God, because there is an escalation in healthcare costs that laboratory medicine could help or could hinder which is not sustainable. In fact, consumers are increasingly interested in quality. That being said, there is no free lunch. All of this will take a lot of work.

Fortunately, there is free literature about standards, literature written, usually by dark poets, often poets who died young like Dylan or Plath. I will let her have the final word.

"Cold worlds shake from the oar.

"The spirit of blackness is in us, it is in the fishes.

"A snag is lifting a valedictory, pale hand;

"Stars open among the lilies,

"Are you not blinded by such expressionless sirens?

"This is the silence of astounded souls."

This is the path forward for standards. Thank you.

[Applause.]

DR. TEUTSCH: That last slide is going to give us a lot to think about.

I'm not sure where to go. I guess we will go to Dr. Cossman.

DR. COSSMAN: That is a tough act to follow.

DR. TEUTSCH: Thank you for being here and talking about a little bit about the clinical perspective from the Critical Path Institute.

Clinical Perspective

Jeff Cossman, M.D.

[PowerPoint presentation.]

DR. COSSMAN: Thank you very much. Steve Gutman is a tough act to follow. But, Steve, I just want to say thank you for all your service at FDA. It has been a real pleasure working with you, and I look forward to whatever you are doing in the future and maybe having a chance to

work with you that way, too.

I'm here to talk to you today about something that we are doing at the Critical Path Institute which may impact standardization of diagnostics in genetics. Let me explain as we go along here what this concept is.

In the development of diagnostics, we can expect delays not just because FDA regulates it but delays in many of the regulatory paths of diagnostics. Many times we see surprises. A diagnostic manufacturer may submit an application to FDA and it may be returned saying, you need to do this again, the data is not prepared in a way that we need, we don't understand it, and you need to redo this for a variety of reasons.

Or there may be surprises on the part of FDA, receiving data that they say is inconsistent or shoddy or not the way that they needed it in the first place.

In order to reduce surprises from either side, we have started to create a standards method that might help both the diagnostic manufacturers and the FDA communicate with each other.

What is needed for this change. This is something that has been a pattern that we have used through

Critical Path Institute. We are a nonprofit agency that is not part of the FDA, not part of industry, and in fact is not part of the government at all. It is a neutral party that helps in communication between the FDA, industry, patient advocacy groups, and researchers in order to communicate among them around science; to improve the methods that are used to develop drugs and diagnostics and bring them to the public and to the consumers.

We have a number of consortia at the Critical Path Institute, or C-Path, which involve multiple companies signing agreements and working with FDA, and in some cases EMEA in Europe, to create best-of-class methods. These can be in safety; efficacy; in the case of Warfarin, dosing; and in the case of Alzheimer's disease and Parkinson's disease, a coalition against major disease in which the largest pharmaceutical companies in the world sign an agreement to work and share data.

What we are talking about here in all of these cases is a way of verifying the quality and accuracy of biomarkers; sharing information across these groups; finding out what is the best-of-class method for predicting safety or efficacy in a particular condition and sharing

that information; agreeing on a consensus on what is the best-of-class method; and having FDA accept this method so that when a company comes with a new submission they will know that the FDA already understands these biomarkers and has, in a sense, preaccepted them as part of their application for a new drug.

Now, what we have seen in running these consortia, because C-Path creates and leads these consortia, is a common theme of diagnostics that are needed. What we felt was there may be a role here for establishing an entity that could provide a means for standardizing the testing of diagnostics before they are submitted to the FDA.

We see many bottlenecks along the way. There are problems in the development of the data that goes to the FDA and the creation, as you have heard, of standard samples. Ten companies may have an assay against, say, troponin or d-dimers, but they are not testing them against the same standard analyte sample. So the data that is coming in to FDA may not necessarily be comparable. So if you are looking for a me-too device or a 510(k), we can't always prove that the test is equivalent because it hasn't

been tested on the same clinical material.

What we are trying to do is reduce the number of surprises that FDA is giving to industry, telling them to redo the study, or the other way around, surprises to FDA from industry. We want to look at ways to improve the efficiency of the requirement for the highest standard of approval at FDA, which is the PMA, and how companies can improve their efficiency in getting to that very high bar.

Finally, there are bottlenecks, as you have just heard, in lack of evidence for payers. How does a payer know whether the test performs as required. An insurance company or CMS is going to pay for a test. What evidence does it have that that test is valuable and actually does the performance that it claims that it does.

So, how do we improve. We improve by the ways that we have already done in the other consortia that we are involved in, and that is to find the best-of-class methods, to look for real proof and real evidence of reliability, and also for a standard submission process. In other words, multiple companies submitting data now submit them in different formats, different kinds of data, different ways of analyzing the data, different clinical

samples. Why don't we standardize that and make life easier for those reviewers at FDA who are looking at diagnostic device applications.

So what we thought was, what we don't have for diagnostics is an underwriter's lab. This would be not a proficiency testing agency like CAP but, instead, further upstream in the pipeline. Diagnostic manufacturers develop tests, submit those for beta testing, say at universities, and that data goes into the submission to FDA.

Why not have a standardized format, a single agency whose sole focus is only on evaluating these diagnostic tests before they are submitted to FDA. They can be an independent body and put a seal of approval on it saying, yes, this test did perform as claimed. We ran it exactly the way it says in the manufacturer's instructions. We ran it on standardized samples. We can attest that, with no incentive as to whether this test is approved or not, it did perform as claimed.

Why not do this in diagnostics. It is done in many other industries: in semiconductors, in food safety, for drugs. This is not a new idea. It is just a new idea for this particular industry.

To quote a famous poet, Steve Gutman, we see that the FDA is interested in this. You have just heard him say the FDA is interested in finding standards for diagnostics. In this case he is talking about targeted therapy. Our original plan was to focus specifically on targeted therapy in cancer, but for this standards laboratory we have heard from industry that they would like to see this service applied and be available for any kind of clinical laboratory diagnostic.

So what Steve told us, as you can see in the middle paragraph, this could be "a template for the validation of diagnostics in targeted cancer therapy," but any kind of therapy. This could be a template and a way to evaluate diagnostics before they go to FDA.

The concept here is to have two levels of evaluation of a diagnostic. One is simply performance. Does it tell you the correct level of whatever the analyte is.

Second would be a much more complex one, and that would be where you have outcomes information attached to the clinical samples so that you could determine the relative value of this diagnostic in predicting a clinical

value such as response to therapy and association with a particular clinical condition.

That information would be put into a report, certified as to the accuracy of the test, and then that data could be used voluntarily by the manufacturer in their submission for FDA approval.

So, what needs does this type of testing meet. One of the goals here is something that this session is all about: having a standard repository of samples that could be used and normalized, and to create methods so that they could be reused as consumed. Then tests could be analyzed on the same samples repeatedly and competing tests could be compared if manufacturers wished to.

It would be a neutral site. It could determine whether or not a new test equals the predicate, or is equivalent to it. For lab-developed tests such as genetics, which may not end up being submitted to the FDA as an in vitro diagnostic, it could be used to evaluate those as well so that providers, consumers, payers, and investors would know whether or not the genetic test or other laboratory-developed test performed as claimed. In other words, did it detect the SNP. Did it do what it was

said to do.

What does this do. It improves reporting to FDA, hopefully improving for the diagnostic manufacturer their chances of having their data accepted. Second, it does provide a format for comparing competing products. If companies wished to, they could have their assays run in a bake-off. You could have multiple companies competing with the same assay, all tested at a neutral site on the same analytes.

All of this information, whether it is competing or whether it is single case-by-case information, provides evidence to the community that needs to know whether or not a test performs as is claimed.

Now, we have talked about this. We are now starting to develop this laboratory. We have seed funding. It is starting in the State of Arizona. The state has provided an economic development package. We have a couple of people who are helping to start this here today with us: Mary Ellen Demars and Ralph Martel. We are looking to take on our first demonstration case, whether it is in genetics or in cancer. We are not sure yet. We are looking for ideas that would fit very specific criteria for first

demonstration cases.

Because people have heard about this, we have been asked a number of questions. One, is this just another regulatory hurdle, which is exactly what I would think this is. I used to run a clinical laboratory. If I had heard about this and didn't quite understand it, I would think the last thing I need is somebody else coming into my laboratory to inspect it and regulate it and find something else wrong.

This is not what this is about. This is not a regulatory body. It has no regulatory authority. It is completely voluntary. The whole idea is to be helpful to the manufacturer or the developer of the diagnostic.

How does this United States Diagnostic Standards Lab, USDS, relate to federal agencies and other agencies that are involved. We are looking at ways of becoming synergistic and complementary. We have had detailed discussions with NIST and Mike Amos as to how they could develop standards for the platforms for this particular testing, as well as with many of the other agencies across federal government.

What happens if the test result comes out and it

is not acceptable or not useful to the manufacturer? They don't have to use it. They own that data. It is their data. They can keep it. It is not published. They can do whatever they want with it. If they don't want to use it, they don't have to use it. They will pay for it. They will be running a fee for service and they can have the data, but if they don't want to use it, they don't have to.

How is IP protected? Everything that is run is confidential within this standards laboratory. If there is any kind of intellectual property or special methods that are being run, those will not be revealed unless the manufacturer wants it to.

How will reference standards be maintained? You have heard methods that are used for that. We know that we need to do that on a case-by-case basis as we enter into this space.

That is the story. I thank you very much for listening and for your attention. Thank you.

DR. TEUTSCH: Great. Thank you.

[Applause.]

DR. TEUTSCH: Why don't we take a couple questions at this point before we move to our final

presentation. Marc.

Question-and-Answer Session

DR. WILLIAMS: This is for Jeff and relates to the last slide. We have certainly seen in other circumstances where "voluntary" things have become ersatz regulatory issues. Look at the NCQA, the Joint Commission, and others. In some sense, if you tie this to data that will be used by payers and other reimbursers, the people that control the purse strings, they may say, we are not going to reimburse any tests that haven't gone through this process. Then you have a de facto regulatory system.

While I think this is really important and this is definitely the direction that things need to be going, I would ask you to respond to that issue.

DR. COSSMAN: I don't know if everybody heard the question. Maybe I can paraphrase it. This could end up becoming too successful in the sense that even though it is not a regulatory body and there is no federal mandate that you have to go through this, it still may be something that everybody wants because the reimbursers, the payers, may require this certification or this process before they pay.

It would then become a de facto regulatory body.

That is a real problem. I can't tell you I have a glib answer how to solve something like that. What we would like to do is start very small with single bites and look at one area and see the pattern that emerges in terms of the reflex of the payers.

First of all, we have to start small because there is no way that you could start with all diagnostics all at once. You are looking at the entire agency so far. We are 2.5 FTEs.

[Laughter.]

DR. COSSMAN: So it is going to be hard to handle all of diagnostics right when we open the door. We are looking for one. One of the criteria would be that exact issue. We have heard that same question from others, that we would be swamped and wouldn't have the bandwidth to be able to manage this and it would become a second FDA. We don't want to be a second FDA. We have no interest in doing that. If that becomes a non-starter, then this won't happen.

But we think that this is so valuable to do, from what we have been hearing from people, that we need to find

a solution to that. I'm open to people who have ideas and are creative and innovative here. We need to be problem-solving. But we don't want to create more of a problem than already exists.

DR. TEUTSCH: Andrea.

DR. FERREIRA-GONZALEZ: To take the next step on that question of becoming a de facto regulator, how do you envision not going that route? What I see is that people start using it and third-party payers get hold of this information. Then you can require an academic laboratory or any other laboratory to send the data to this place in order to be reimbursed by any of the third-party payers.

DR. COSSMAN: I think it is a similar issue. How do we not become a regulatory body. That is in terms of payers. Is that what you are asking? If payers would require it, then you would become a de facto regulatory body. I think it is a similar point.

We don't have the solution for that. What we are saying is we would start small, with a single example, move out from there, and see what emerges in terms of the pattern from payers. We are just starting our discussions with payers to see how they would react to this.

In fact, the very first one I talked to -- and I won't say what company, but it is a very large insurance company -- said, we at the insurance company don't have the bandwidth to be able to determine which test someone ran. We just pay a CPT code. We don't know if they ran the test that worked well or the test that worked medium well or the test that doesn't work at all. We don't have an inspection method to be able to determine that. So right now, they wouldn't even be able to use this information. Even that hasn't happened yet.

DR. FERREIRA-GONZALEZ: They don't have the means today of identifying this, but they can ask that. If you are going to be submitting claims to particular third-party payers, then you submit information that you have been cleared.

DR. COSSMAN: They could.

DR. FERREIRA-GONZALEZ: We already have regulatory bodies to look at the quality of the testing. It seems to me that it could be, in the future, another hurdle to this.

DR. COSSMAN: Exactly. If this looks like it is an insoluble problem and is another hurdle, that is a deal-

stopper. What we want to do is be innovative and creative here and find solutions for getting through this so that we can find ways around it. I don't have the answer here today, but if people have ideas, we are open to suggestions. I would be happy to talk to people in the insurance industry and CMS and see if there are ways that we can do this so that it works in a way that doesn't open up a floodgate of problems but rather is problem-solving.

DR. TEUTSCH: Great. I know we would like to have some more discussion. Thank you very much, Dr. Cossman. We appreciate that and your initiative in addressing this important topic.

Our final speaker is Mike Amos, who we all know. He will talk a little bit about the future directions in clinical diagnostic standards development.

Mike, we are going to hold you to your 10 minutes so we do have time for some discussion at the end. Take it away.

Future Directions in Clinical Diagnostic Standards Development

Michael Amos, Ph.D.

[PowerPoint presentation.]

DR. AMOS: Not a problem, not a problem. Thanks for your attention. I hope you appreciate the level of detail and precision that my NIST colleagues go to to provide standards for the various applications. I think John's table that talked about the various levels of who uses them and then Dave's table talking about the horizontal versus vertical standards gave you an idea about how we think about things.

I should probably bring my other hat up here because my boss, who is Dr. May, told me to put this disclaimer on here. I'm going to talk about things that we have learned over the last couple of years through many talks with many different people about what they consider the future of diagnostics and where things are going. At the same time, these are not official NIST programs or ideas but just food for thought for you.

What I want to talk about today are some of the harsh realities that are really going to drive health care change in the future, some lessons learned and what I think will happen, the fact that laboratory medicine will drive a lot of this change, some measurement challenges and the

role measurement technologies and standards will play, and a potential plan to enable the change.

Where we are is kind of scary when you consider that about 83 percent of our total health care costs go to cover chronic diseases, whereas the rest of it is only about 17 percent. This constitutes almost $1.7 trillion out of the $2 trillion that we spent in 2005. Forty-three percent of that is spent on hospitalizations. The scary part is the most expensive to treat are among the fastest-growing reasons for hospitalizations, according to AHRQ.

Millions of people suffer from diseases that there is little known about the genetic basis. We have a growing number of problems with kids taking drugs for chronic diseases. More and more kids are being diagnosed with chronic diseases for which they are being treated. Diabetes is running rampant and growing at a rate of about, I think, 5 percent a year for type 1 diabetes. Kids under the age of five are now taking drugs for type 2 diabetes.

The problem is that things are not going that well in medical research. The innovation gap is really widening. There is more money going into research with not great returns on investment. There are more and more

manufacturer-reported adverse events to the FDA all the time. It has grown dramatically since 1990, with billions of dollars of drugs coming off the market because of toxicity.

The future is not that great for diagnostics, really, if you base it on what has happened since 1995. This is, as best as we can tell -- and Steve's group helped me put this together -- the complete list of single protein biomarkers that have been approved by the FDA. There may be one or two recent ones. But I went through the FDA website again before I did this, and I couldn't find any more.

So things are not really looking that great in the future. Our grandchildren are going to be spending more money than they earn on health care. Like Steve said, these trends are not sustainable and a new development paradigm is really needed.

So, what have we learned. We have learned that the human body is very complex. It is really not just made up of all those individual components. Really, disease is caused by perturbations in very, very complex biological networks. It is not simple pathways anymore. Forget what

you learned in high school. There is no such thing as a metabolic pathway. It is one of these globby things.

So, what have we learned. Disease is a result of perturbations in these pathways. Genomics has been helpful, and it will continue to be helpful but it is limited. Only a very small number of single protein biomarkers are good indicators or predictors of a limited number of diseases, and more complete understanding of human physiology is needed in order to identify good biomarkers.

What is going to happen. Medicine will focus on keeping people well. It has to. The only way we are going to really catch up in health care is by keeping people out of the hospital. That is possible. The way to do it is the fact that laboratory medicine will probably lead the way. -Omics will dominate. Complex disease signatures that are comprised of hundreds or thousands of data points will really be the biomarkers of the future.

Drug companies will develop their markets around interventional therapeutics and treatments like cholesterol and statins. They will use the same model. It will be based around these complex disease signatures. Disease

signatures are measurable alterations in complex biochemical networks.

So, what happens. You get abnormalities in all this stuff, and you can do multiplex measurements and computer integration to develop disease signatures. There are a bunch of these things. We have no idea what these disease signatures are going to look like. Probably, it is going to be some sort of risk score, a number from one to 100, whether somebody is going to get this disease or not, but we really don't know what that is going to be. We hope that it is going to enable scientists and physicians to make better decisions.

Discovery decisions will increase the drug pipeline and all those things. Better clinical decisions help people, not just the drug and diagnostic companies.

Really, in between wellness and symptoms are these transitional states. That is where the focus is going to have to be. We are really looking at markers that occur years before disease symptoms occur. They often occur long before people realize they are sick.

They are unique biochemical markers. They can distinguish health from sickness. They are going to be

person-specific. The rules of clinical trials are going to have to change because each person will end up serving as their own control.

There are typically going to be parameters in blood. Those probably are the true biomarkers that we are all looking for and that could be detected with proper technology.

A disease signature is like a radar signature. A good radar operator can identify a blip on a radar screen that is a bad guy versus a good guy. What we want to be able to do is develop similar technologies in the future for diagnostics.

One potential concept is being espoused by Dr. Lee Hood, who talks about organ-specific blood protein fingerprints as a potential way to do this. He calls it systems medicine. It integrates measurements and computers. It is basically taking a drop of blood, putting it on some analytical platform, putting it in an instrument, and then getting some data out to enable the complete visualization of what is going on in your body. That is the dream.

Why is this critical and what is going to happen.

Today the healthcare markets are based on the number of sick people. Every drug company bases their market numbers and projections on the number of people they can treat. That is based on the number of people that they project will come down with a disease based on historical data.

The metrics of morbidity and mortality show the outcome is that people suffer and die of chronic diseases. It is not changing. We will see $4 trillion in healthcare costs projected by the year 2015. Like Janet Woodcock said, that is probably not sustainable.

The healthcare markets could be based on the number of people with preventable diseases. If that were the case, the metric would be the number of people positive for a valid predictive biomarker. The outcome would be that more people would die of trauma and in their sleep from old age, rather than spend 70 percent of healthcare dollars in the last two years of their life in terminal care.

Potential savings are, just for diabetes, probably at least $50 billion. Diabetes is more expensive to treat than cancer. We all know that.

What is going to happen is visualization of

disease signatures. What kind of standards will be needed for this type of thing. We are really talking about the complete spectrum, but we will have to take a very logical and structured approach to it and take into account all the things you heard today from my colleagues: horizontal versus vertical standards, and what are the highest priorities of things that we should go after.

That is really what Willie talked about. We felt, and the community felt, that protein measurement science is probably one of the biggest challenges.

These are some of the things that we are going to have to do. But two fronts are really to promote discovery of disease signatures and then, on the back end, clinical analysis of these disease signatures.

I love my boss, but I have to disagree with you. We will always have this conversation, Willie. I think, coming from industry, if I had had a set of standards that I could anchor my tests against where I didn't have to guess and empirically try to figure out what my assays were really doing, then I could have sped up things a lot in my assay development.

I think the things that Dave is trying to do with

proteomics and anchoring what I call the platform standards of mass spec to make sure that your mass spec works properly, are going to really drive the future.

You have transition states and systems medicine. That is one approach. Developing disease signatures to usher in the age of individual therapeutics and improve quality of life and help in economic security, which is, as Willie showed, part of our mission.

What is preventing us from getting there. Basically, it is the capabilities of doing these things, among many other things, but these are pretty much the major issues. It is really doing these types of measurements and the ability to analyze these types of things.

Here is a potential opportunity and a potential way of stimulating the advent of new technology. I think we are woefully deficient in our ability to measure proteins, and that is a real issue. I think we are at about the same place we were at the beginning of the Human Genome Project.

One way to stimulate interest is to have a mission to the Moon. So here is an idea. Maybe we can put

a stake in the ground and say we can identify disease signatures for the most important diseases by the year 2020. The number is obviously subject to debate, but these are the kinds of things that we would have to do and hopefully will enable some new approaches and a better way of looking at diseases and keeping people healthy.

What do we hope to learn? We have some pretty lofty goals here, but I think without new technology it is not going to happen.

One thing I can say is, when I came to NIST I was pretty ignorant of all this. I hope that the presentations today really helped you get an appreciation for what my colleagues do. I am amongst egghead scientists who focus on the nitty-gritty, nuts and bolts of measurement, and I think that that is why we are here. I appreciate your attention.

DR. TEUTSCH: Thanks, Mike.

[Applause.]

DR. TEUTSCH: Thanks to all of our speakers. We have obviously had a tour from the importance of getting measurement accurately to what the future world might look like.

We have just a few minutes, and I think we should take this opportunity to ask questions of any of our speakers who are still here or to have a discussion among ourselves. Let me open the floor for a couple of questions.

Discussion

DR. TEUTSCH: Let me ask you, do you have any additional comments that you would like to make from the CDC perspective?

DR. KALMAN: We think that having reference material is really key to assuring the quality of these tests not only for the day-to-day QC of the tests but also for proficiency testing, which is a big deal. It was quite a large part of the Oversight report that this group did a few months back.

We did a count. I think there are about six different diseases for which there are higher-order reference materials either from NIST or FDA or something like that. We count six. On the Gene Test website, there are over 1,300 genetic tests currently available. That is a really small fraction of the current tests that are available.

So the CDC, through the GeTRM program, is trying to address this gap by just simply organizing a volunteer effort among the people in the genetic community. We are just characterizing publicly available cell lines and DNA from the Coriell repository so that we have a larger supply of materials so that we can feel confident in knowing the genotype of these and so labs can use them for quality control and also the proficiency testing needs.

Right now the projects that we are working on are pretty much all being driven by requests from CAP for proficiency testing materials. We are starting a real large project for pharmacogenetic materials. We are going to do over 100 DNA samples for five pharmacogenetic loci. We are going to get other data from other labs as well on other loci. We are going to try to do a project for array CGH.

We were trying to do a project for Duchenne muscular dystrophy, which is something that CAP asked me to work on, but all the labs are stopping their testing because of the patent issue. So I don't know what is going to happen.

DR. TEUTSCH: Coming full circle. Andrea.

DR. FERREIRA-GONZALEZ: I want to thank Lisa for a tremendous effort and the role that she has played at CDC in getting the GeTRM program started and being one of the strongest advocates for this. I think she needs a round of applause from all of us.

[Applause.]

DR. FERREIRA-GONZALEZ: That said, like you said, there is a lot more work that needs to be done. But I think it is interesting that you have already identified through the collaboration with professional organizations or end users of different laboratories what are the current needs of the laboratory not only in proficiency testing but also reference materials that we can use to analytically validate the assays and continue quality control.

I was wondering, what is the level of cooperation between the GeTRM program and the NIST genomic program. I think a lot of the work that you have done in identifying some of the needs can be translated and the deployment of the work NIST can take over.

DR. KALMAN: I do talk to NIST on a regular basis. Our program has a yearly advisory committee meeting. We always have a few people from NIST at our

meeting, so I talk to them. Also, in the area of molecular oncology there are a few people from NIST that I have been talking to.

So, yes, I try to keep the communication lines open. But if you want to talk some more, that would be great.

DR. BUTLER: Margaret Klein went to the meeting that you had last month. We are looking forward to working more with you in the future as we get more into future genetic tests.

DR. TEUTSCH: Marc.

DR. WILLIAMS: I was going to ask Andrea's question. But then as Mike spoke, I said, if that is the vision of where things are going, then in some sense is investing a lot in genomic validated samples really worth it if we are really going there.

I guess the question that I have -- and probably you or Dr. May would be the best ones to address it -- would be, what is your real vision about where you are going to need to invest your limited funds in terms of standards in the biomedical realm? Is it going to focus on genomics? Is it going to focus on proteomics or

metabolomics? Are you going to try and do it all?

DR. MAY: I think, in the short term, Mike's vision is 2020. We have a lot of living to do between now and then.

Certainly, in the short term, the focus of the NIST's new activities is going to be on medical imaging and protein measurement science, for sure. Beyond that, we might do some other things.

If you are looking at the near future, I think for the next two to five years the emphasis is going to be on improving our capabilities to support medical imaging and developing more core competencies in protein measurement science.

That would address lots of things. It would address this disease signature issue that Mike talked about, as well as the issue of follow-on biologies.

So we are trying to increase our core competencies and put more tools in the toolkit to address a number of things. Now, in the longer term, we are still going to continue our work in genetics. We are not going to stop those things. But if you look for areas that across all of NIST we are going to expand in, it would be

those two.

Now, putting on my director of the Chemical Science and Technology Laboratory hat, certainly in the Biochemical Science Division there is going to be a greater emphasis on genetic testing and DNA-based diagnostics. As John mentioned to you, we have just done some reorganization within our Biochemical Science Division to address just that issue.

DR. WILLIAMS: In follow-up to that, our Oversight report identified, as Andrea pointed out, that this PT issue and having samples is a huge issue. We have 5,000, plus or minus, genetic tests that are out there and a small fraction of those actually have PT materials that are available and in use.

From what I'm hearing you say, I think it may be unrealistic to expect that NIST is going to be the savior riding in on the stallion at this point.

DR. MAY: That is true. But certainly, if that is a major issue that your Committee has identified, sending a note to me to that effect, perhaps with a copy to the acting NIST director, would not be a bad idea.

DR. AMOS: Marc, just let me say one thing. It

is clear that genomics is going to be an integral part of the disease signature. I think that the discovering technologies of the future are really going to focus on the ability to understand the environmental effect on the genome. So you have to have good genomic data to do that. There are all sorts of issues with the sequencing things that are going forward.

I think my colleagues have decided that genome-wide association studies are something that we don't want to do. We are looking at next-generation sequencing. I will put it that way.

DR. TEUTSCH: Mara, you get the last word.

DR. ASPINALL: I think I also, once again, agree with where Marc is going. So this has truly been a red-letter day.

DR. TEUTSCH: It is a great place to end the meeting.

DR. WILLIAMS: She is going to hit me up for a drink later.

[Laughter.]

DR. ASPINALL: The question really, Steve, was to you. I think this was a great session, with the ability to

hear the different perspectives of what is happening today and getting the various approaches to that. What role do you see SACGHS taking? This is great information, but I know that tomorrow we are going to jump into priorities going forward. Where do you see this going?

I love the idea of taking some action and sending some letters to NIST. As Marc said, this is, to me, entirely consistent with the recommendations not just in the last report but in the last two that talk about gaps and the need for essentially standard-setting or ensuring quality across the system. Now we have an opportunity that doesn't require potentially major changes in legislation by Congress or otherwise but just a prioritization. I would vote for taking some action to at least enforce that.