FDA PUBLIC MEETING

CLINICAL ACCURACY REQUIREMENTS FOR

POINT OF CARE BLOOD GLUCOSE METERS

March 16, 2010

DR. HARPER: Good morning everyone. My name is Courtney Harper and I'm the director of the Division of Chemistry and Toxicology Devices at the Food and Drug Administration, and I'd like to welcome you all to our public meeting on blood glucose meters. We're very excited to have you all here, and we really think that this is a great opportunity for us at FDA to hear from stakeholders about this important topic, blood glucose meters and their use in the lay-user population, and also in health care facilities and settings.

So I hope you all are looking forward to the next two days as much as I am. I'm going to give a few logistics, and then I'm going to introduce Dr. Jeff Shuren, our Center Director, who is going to officially open this meeting.

So a couple of housekeeping items. For lunches on today and tomorrow, lunch is on your own. The hotel actually offers a buffet and other menu items, but we've also provided in your packet some information about some of the local lunch areas that you can find nearby, within walking distance or

driving distance.

We will be strictly adhering to the start times of the sessions, so please be prompt in returning to your seats. I believe there's a bell that may significant the start of each session. And we ask that you turn off all cell phones to be courteous to those around you.

So before I review the format for today, I am going to remind you that transcripts of this meeting can actually be obtained approximately ten days after the meeting, and the instructions for obtaining those transcripts are actually located in the FR notice, which can be linked through this public meetings website. So you can go on line and find out how to get transcripts of the meeting if you're interested. We ask that you actually contact speakers individually if you would like copies of their presentations, to ask them if they're willing to share them.

So the format of this meeting, we actually have three sessions; two sessions are today. The first session is on the clinical accuracy requirements

for blood glucose meters, and the second session is on glucose meter interferences and limitations. And tomorrow we will be talking about tight glycemic control in health care facilities, and also a couple of topics on user error and liability.

For each of these sessions the moderators will be introducing the topic. If there is time at the end of each presentation, we'll take a few questions. Then there will be presentations from the speakers in that area, and following that there will be a panel discussion by panelists. The panels will consist of the speakers from that session, and in some cases a few other selected experts.

Our moderators will lead the panel discussions. The moderators may have some questions for the panel that they want the panel to address, but we would also really encourage audience participation in that portion of the program. We would really like people to come up to the microphones that are set up in the aisles, and address some questions to the panel that you would like discussed. Also, if you have any comments or points of view, we would be very happy to

hear from you on those.

The microphones are set up in the aisles, as I said. We ask a couple of things, though; first, that you identify yourself prior to asking a question or giving a brief comment, and second, that you limit your comments to a minute or less so that they can be brief. So questions to the panel or brief comments or a point of view are quite welcome.

So now it's my great pleasure to introduce Dr. Shuren. Dr. Shuren became the Director of The Center for Devices and Radiological Health in January of 2010. Before that he was our acting Center Director since September. Our Center is responsible for assuring the safety, effectiveness and quality of medical devices, assuring the safety and quality of radiation-emitting products, such as cell phones and microwaves, and for fostering device innovation, which is why we're here today.

Dr. Shuren received his Bachelor of Science and Medical Doctorate Degrees from Northwestern University under its Honors Program in Medical Education. He completed his Medical Internship at

Beth Israel Hospital in Boston, his Neurology Residency at Tufts New England Medical Center, and a Fellowship in Behavioral Neurology and Neuropsychology at the University of Florida. He received his J. D. from the University of Michigan. He has held various policy positions and planning positions within the FDA from 1998 to 2009, including Acting Deputy Commissioner for Policy, Planning and Budget, Associate Commissioner for Policy and Planning, Special Counselor to the Principal Deputy Commissioner, Assistant Commissioner for Policy and Medical Officer in the Office of Policy.

Dr. Shuren has served in the leadership role at FDA, or on behalf of the Agency, on numerous initiatives, just to name a few, including the reauthorization of The Medical Device User Fee Act, the creation of the Sentinel Initiative and the development of FDA's Pandemic Influenza Preparedness Strategic Plan. So we're in good hands, and I would really like you to join me in welcoming Dr. Shuren.

DR. SHUREN: Thank you, Courtney. Good morning. Can everyone hear me in the back? Good. I

ask because I'm known to mumble, and this actually got me in trouble the other week. So I don't know if you've ever been out to our White Oak facilities, but if you go, particularly the conference rooms on the ground floor, they have glass windows. And just about ten days ago I had this professor come in, an elderly gentleman, very nice, and he started talking and I started to answer his questions. He goes, well, I can't hear you, I'm a little hard of hearing. So I speak up. I really can't hear you. So the next thing I know, I'm talking very loudly. After the meeting some people came up to me and said, why were you yelling at that old man? So I want to make sure I get it right.

Well, thank you for joining us, either in person or on webcast, for today's meeting on the clinical accuracy of blood glucose meters. If you're viewing via the web and would like to comment on anything that you see or hear today, I encourage you to submit written comments to the docket, which will remain open until April 20.

Diabetes affects individuals of all ages,

from infants to the elderly. There are currently at least 24 million Americans with either Type I or Type II diabetes, and likely many more cases that have yet to be diagnosed. Over the past 20 years, the number of people diagnosed with diabetes has increased from 30 million to 250 million worldwide. And the prevalence of the disease, if it continues to increase, is estimated to be, in 2025, as many as 350 million diabetics worldwide. The numbers are staggering.

Without proper management, diabetes can have devastating consequences. Let me quickly run through some statistics. Diabetes increases the risk of cardiovascular disease by up to a factor of four. It accounts for thousands of emergency room visits each year, and is the leading cause of kidney failure and adult onset blindness. Additionally, complications for diabetes can lead to more than 80,000 amputations each year.

The importance of blood glucose meetings in the management of control of diabetes is unquestionable. Thirty percent of those diagnosed with diabetes require insulin, and are likely to use blood

glucose meters. These devices are used by patients themselves for self- monitoring, and also by health care providers in a variety of clinical settings such as hospitals, emergency response units, nursing homes and physicians' offices.

We're here today to discuss issues related to point- of-care use of blood glucose meters. FDA receives approximately 12,000 adverse event reports associated with blood glucose meters each year. Many reports highlight issues related to the analytical and physiological limitations of these devices to the lay users, whether at home or in the clinical setting, influence the performance of these devices, and to the way performance may vary under different conditions of use. This workshop will focus on three important topics related to blood glucose meters performance.

Session 1 will focus on the clinical need for accuracy in blood glucose meters, and the reality of what point-of-care meters are capable of achieving. To evaluate pre-market submissions of blood glucose meters, FDA currently applies the principles outlined in our own guidance on glucose meters, and ISO 15197

Standard. Some in the clinical and patient communities have questioned whether the current FDA recognized accuracy standards for blood glucose meters are acceptable, and have challenged FDA to require tighter performance requirements. Others believe the current analytical performance of glucose meters is adequate, and that there is no evidence to support the need for higher standards.

We are interested in hearing about what the appropriate analytical and clinical accuracy requirements for blood glucose meters should be. This discussion is very timely, since we are currently updating our guidance and the ISO Standard is going through revision.

Session 2 will focus on medications and other substances that interfere with the technologies blood glucose meters employ. Performance of blood glucose meters may be affected by administered drugs, common physiological conditions such as diabetic ketoacidosis and user interface issues. For example, administration of therapies containing maltose, which are commonly prescribed to patients in the hospital,

have resulted in falsely elevated glucose readings. We are interested in hearing about analytical interferences and physiological limitations, an issue that is not unique to blood glucose meters, but rather one that the majority of point-of-care technologies face.

The third and final session, to take place tomorrow, will focus on the use of blood glucose meters for tight glycemic control in clinical settings. Despite the fact that these devices have not been approved for this use, glucose meters are increasingly being used to achieve tight glycemic control. Over the past three decades, blood glucose meters have become smaller, faster and more accurate, and they now allow for better glycemic control than in the past. However, there is no consensus that blood glucose meters currently on the market are accurate enough to be used in this way. We're interested in hearing about the benefits and risks of using glucose meters to achieve and maintain tight glycemic control.

We've assembled an impressive and distinguished list of speakers for this meeting. They

will provide a basic framework from which to begin our discussions. The roughly 400 participants and attendants at this workshop, you, form an equally impressive group. You represent health care providers, diabetic educators, professional organizations, and perhaps most importantly, the people who use these meters every day, either for self- monitoring or for clinical care. I encourage you to engage in the discussions today and tomorrow, share your questions, as they might shed further light at the issues at hand, and please also share your ideas for solving the problems we will be discussing today and tomorrow.

I'll conclude by leaving you with two critical questions that we want you to keep in mind and address during this workshop. First, how should the FDA address and balance the clinical needs of diabetics and the technological limitations that are inherent to fast and relatively simple blood glucose meters. Second, what steps should the FDA take to improve the quality of point-of-care blood glucose meters, and what are the responsibilities of industry,

the health care professional community and consumers, to assure that point-of-care blood glucose meters are safe and used safely.

Thank you very much for joining us today and tomorrow, and I look forward to a very lively dialog.

DR. HARPER: Thank you, Jeff. I'd like to now introduce Dr. Bill Clarke. Dr. Clarke is the Robert Blizzard Professor of Pediatrics and Chief of Pediatric Endocrine Division at the University of Virginia School of Medicine, in Charlottesville, Virginia. He is the author of over 130 Journal articles, has served on several editorial boards, and has been associate editor of Growth, Genetics and Hormones for the past 15 years.

Dr. Clarke is a graduate of Duke University and Vanderbilt University School of Medicine, and did his pediatric and endocrine training at Washington University in St. Louis, Missouri. He is board certified in Pediatrics and Pediatric Oncology.

You all are familiar with Dr. Clarke's error grid, which has been useful in the evaluation of the performance of glucose meters. His research

interests involve understanding glucose counter-regulation and devising methods for improving glucose control in children with Type I diabetes, including recent contributions regarding the accuracy of continuous glucose sensors and the artificial pancreas. Dr. Clarke has been honored by the American Diabetes Association for his work and outstanding contributions to diabetes in youth, and has been listed in Best Doctors in America. Welcome, Dr. Clarke.

DR. CLARKE: Thank you. I'm always a little bit technologically impaired. Well, I was given ten minutes to be the moderator, so I am going to take my ten minutes but no longer.

I do have some questions and some comments that I think are important for us to think about this morning as we talk about clinical accuracy requirements for blood glucose meters, the first of which is what we mean by accuracy standards for glucose monitors. This is the current FDA requirements for blood glucose monitors, for self-blood glucose monitor accuracy, and I think that we're

going to be spending a lot of time talking about ISO criteria, and maybe even some time talking about the error grid analysis and strictly clinical accuracy of the systems a little bit later.

Now, from my standpoint, clinical accuracy means that the information that is presented to one can result in a clinically accurate treatment decision. And those decisions may be different in different situations, and I think we're going to talk about that over the course of the next couple of days. Specifically, self-blood glucose monitors were developed for patients and for health care professionals to guide clinical decisions. And there are a lot of us who were around at the time when they were developed, so that people didn't have to collect urine samples to make clinical decisions. With these systems, I think it's important to point out that we conducted a 10-year diabetes control and complications trial, which is the largest clinical trial ever, and we got very, very significant results. So they can really, really help us, even in their current form.

The other important factor is, I am unaware

of any superiority studies with one meter versus another. Do people have less hypoglycemia on one system than another, more DKA, do they have better hemoglobin A1Cs?

Fourth, here, it's very important that the ADA has said that they are not acceptable for diagnosing diabetes, and the FDA will tell you that their use is really not approved in intensive care units.

So, where are the problems and what are the problems that we're kind of looking for? I think it's going to be a little bit of a challenge for us to look at this, but here are some reasons I think of for concerns regarding accuracy: Hypoglycemia, it's still the barrier to normal blood sugar levels. Glycemic variability causes oxidative stress. Failure to achieve hemoglobin A1C targets. Cognitive dysfunction, which is now reported in people with frequent low blood glucose and frequent high blood glucose. Depression, which is rampant in the diabetic population. And new uses, such as in the intensive care unit. Krouwer and Cembrowski, and I saw Dr.

Cembrowski here just a few minutes ago. I didn't speak to him, but they talk about total error, and this is part of the presentation you're going to hear later. But if you look at total error, or total accuracy, which is the way I changed it, notice that analytical, clinical and statistical really only form a small portion of the accuracy, and I think that's because there are limitations. And those limitations may be particularly in the area of interpretation and response.

A paper in diabetes care last month, with an internet survey of people with Type I diabetes, demonstrated that over 50 percent of people omitted insulin, and over 30 percent intentionally omitted insulin. Peter Chase at the Barbara Davis Center has shown that if you omit one injection a week, it's approximately a .5 increase in the hemoglobin A1C level.

So what are we trying to really achieve here? Dr. Breton's going to tell us about some simulation studies later, and insulin delivery. If you're talking about giving a dose of five units or

less, well, we know that the accuracy of that is at least plus or minus one unit, that's 20 percent. For children that's a big deal, because that may be what their insulin dose is prior to a meal. And then there are all kinds of health care professional factors that I'm not going even begin to talk about.

I like to always show the error grid. I think most of you are familiar with it, but I want to also show you how it can be modified when one changes the target range. So here you see the target range is 70 to 180. Here is the target range of 80 to 110, which is what's commonly being talked about for intensive care unit purposes. And notice that what happens is, that when you scrunch that target range, you increase the errors by a significant amount.

Pardon my drawings, but the computer program for doing error grid was down, so I can guarantee you that this is accurate. I'm not a good colorer and actually, my junior faculty associate pleaded with me to be allowed to make these on his computer, and I told him, no, he needed to go do his research.

Anyway, the original error grid is on the

left. Here on the right is what happens to the error grid when you say, well, we're only going to have plus or minus 5 percent be considered accurate. So you really, really squeeze down that accurate range and increase again your other errors, and perhaps your benign errors as well. If you do both, squeeze down the zone A or the accuracy to plus or minus 5 percent, and move the target to 80 to 120, what you see is that there's a very narrow window in which you can be clinically accurate, and lots of overtreatment. Either too much insulin or not, decisions that could be made, at least based on what we have right now.

Finally, I think we need to talk about how we will evaluate a change in meter accuracy. Will superiority be our goal, or will equivalency -- I apologize for the spelling -- be tolerated, and poor spelling can be certainly tolerated. At the current moment, I think there's only one decision that can be made using a blood glucose monitor, and that is, my blood sugar is low and I need to eat something now. And any other decision is going to be based on what the previous glucose is, and so maybe we really need

to be talking about continuous systems and trends and rates of fall and direction of change, in order to make an intelligent clinical decision.

Well that's my last slide. I hope that I've given you some things to think about over the next day or so, and we will move on to our first speaker, who is Patricia Bernhardt. Ms. Bernhardt is a scientific reviewer here at FDA, and she is involved in pre-market clearance and approval, and has been since 1997. And so she is going to present for us the FDA perspective on evaluation of point-of-care blood glucose meters. And would somebody like to clear that, and we'll escape this, maybe, and go to your talk.

MS. BERNHARDT: Thank you, Dr. Clarke, and good morning, everyone. Blood glucose monitoring systems are used by diabetics in the United States every day. They have been around for more than three decades, and during that time they have become smaller, faster, easier to use and more accurate. They are used not only by diabetic patients, but also by health care providers in a variety of settings such as

hospitals, nursing homes, emergency response units and in physicians' offices.

Before I talk about FDAs evaluation of blood glucose monitoring systems, I'm going to provide a brief overview of the medical device regulations. Congress gave FDA the authority to regulate medical devices under the 1976 Medical Device Amendments to the Federal Food, Drug and Cosmetic Act. So, how are medical devices regulated?

They're regulated by intended use, they're risk-based by intended use in to three classifications: Class 1, which is low risk; Class 2, which are moderate risk; and Class 3, which are high risk.

Now, how do blood glucose monitoring systems fit into this picture? Blood glucose monitoring systems are in a category of medical devices called in vitro diagnostic devices, or IVDs. By definition, an IVD is a reagent instrument or system intended for use in the diagnosis of a disease or other conditions, including a determination of the state of health, in order to cure, mitigate, treat or prevent disease, or

its sequellae in man. They are for use in the collection, preparation and examination of specimens from the human body. IVDs are used in clinical laboratories, point-of-care sites such as operating rooms, emergency room, nursing units, nursing homes, and by patients at home.

FDA regulates IVDs by their intended use and the risk of an incorrect result. For example, a high-risk IVD could be an HIV test, where the risk of a false negative result could expose others to the disease. All IVDs must establish adequate analytical and clinical performance, and IVDs have their own unique labeling regulations, which require that certain information such as intended use, limitations and performance are in the product labeling.

So blood glucose monitoring systems as IVDs are Class 2 devices that have a moderate risk. They require FDA clearance. They must be substantially equivalent to a predicate device, which means that they must be at least as good as, but do not have to be better than, a device that is already on the market with a similar intended use. FDA evaluates the

intended use performance and labeling of glucose meters. So a typical intended use for glucose meter is that they are for the quantitative measurement of glucose and whole blood by lay users at home, or by health care professionals in clinical settings, to assist in the ongoing evaluation and management of individuals with diabetes. They are for monitoring, not for diagnosing or screening. And currently there is no distinction between the performance requirements for over-the-counter and professional use glucose meters. So when a glucose meter is cleared for over-the-counter use, it can also be used in professional settings.

The components of a blood glucose monitoring system typically include a meter, test strips, quality control solutions and sometimes lancing devices, lancets and alcohol wipes. Although sometimes more than one meter from a single manufacturer can use the same test strip, FDA considers each meter and test strip, when used together, to be a separate system, and each system requires its own performance and is evaluated separately.

There are different samples types that can be used with glucose meters, and each sample type requires FDA review and clearance. All meters typically use capillary whole blood from finger sticks, but in addition, some use arterial, venous or neonatal blood. Also capillary blood from sites other than the finger, such as the forearm, upper arm, palm, thigh or calf, is sometimes used. This is known as alternative site testing, or AST.

Glucose measurements from alternative sites can differ significantly from measurements from the finger at certain times, and require specific instructions for use that define the appropriate times when alternative site testing can and cannot be used. There are a variety of guidances, guidelines and standards that FDA currently uses in the evaluation of blood glucose monitoring systems. This list shows some of these documents. The documents provide recommendations for the types of information and data to be evaluated for glucose monitors, and are used by industry to prepare FDA submissions, and by FDA when reviewing the submissions.

To evaluate the performance of glucose meters, FDA looks at factors. To evaluate the precision of a blood glucose monitor, we look at repeatability and intermediate precision. We evaluate glucose concentration spread across the measuring range, and determine the variation and repeated results over time. One way to evaluate accuracy is with method comparison. Method comparison studies are conducted to demonstrate how well results compare to a reference method. A reverence method, such as the YSI, is defined as one that is well validated for precision and trueness, and is traceable to a recognized glucose standard, such as the National Institute of Standards and Technology standard reference material.

Method comparison studies are conducted on a minimum of 100 capillary samples that span the measuring range of the device. Because it's hard to find samples with glucose concentrations at the extreme low and high ends of the measuring range, a small number can be spiked or altered to achieve those levels.

The results of the method comparison study are evaluated with a variety of statistical presentations to determine system accuracy. The term, system accuracy, means how well the results agree with truth. And total system accuracy includes elements of imprecision and interferences as well. The current FDA minimal acceptable system accuracy, and accuracy in the hands of lay users, is that 95 percent of individual glucose results shall fall within plus or minus 15 milligrams of the results of the reference measurement at glucose concentrations less than 75, and 95 percent of individual results shall fall within plus or minus 20 percent at glucose concentrations of greater than or equal to 75. This criteria is the current recommended ISO 15197, standard criteria for system accuracy.

Now to see how well the device performance meets the minimum acceptable criteria, the data from the meter is presented in two tables that show the number and percent of results that are within 15, 10 and 5 milligrams per deciliter of the reference results for samples with concentrations less than 75,

and within plus or minus 20, 15, 10 and 5 percent of reference results for samples with concentrations greater than or equal to 75. This format is recommended by ISO 15197. A minimum of 95 percent of the results must meet the minimum acceptable criteria for both less than 75, and greater than or equal to 75.

It's interesting to note that in a recent internal evaluation of glucose meters cleared in the last two years, we saw that approximately 72 percent of them would meet a plus or minus 10 milligrams per deciliter for concentrations less than 75, and approximately 50 percent would meet plus or minus 15 percent at concentration greater than or equal to 75.

In addition to method comparison, we evaluate performance in the hands of lay users. We look at studies conducted with a minimum of 100 participants. The participants collect and measure their own finger stick samples, using only the instructions for use that will be provided with the marketed device. A health care professional also obtains samples from the participants at the same time

to run on the reference device. A statistical comparison is made between the lay user results and the reference, and the results are presented in the ISO recommended tabular format that I just showed you. In addition, FDA evaluates the results from questionnaires given to the study participants to assess the readability of the labeling, and the ease of use of the device.

When a claim is made for alternative site testing, FDA also evaluates data from studies where lay users obtain and run their own samples from each alternative site being claimed. The subject should be in a steady state condition, which means times when their blood glucose is stable, and results from each site are compared to a sample obtained by a health care professional at the same time, and tested on a reference method. This data from each site must meet the minimum acceptable accuracy criteria, and is also presented in the ISO recommended tabular format.

Other sample types that we evaluate are venous and arterial when claims are made for those matrices. Data from study comparing the results from

each matrix to a reference are evaluated. The samples are collected in the appropriate anticoagulant, and they span the measuring range. For neonatal claims, as mentioned on the slide that showed the guidance documents that we use, there is an FDA guidance that provides recommendations on the study, design and data collection for neonatal studies. The critical glucose range for these samples is 10 to 15 milligrams per deciliter, which hematocrits between 45 to 65 percent, so extra attention is paid to samples with those values.

FDA also evaluates linearity which refers to the ability of the device to provide results directly proportional to the true concentration across the measuring range. In other words, how well the device results compared to true results conformed to a straight line. We evaluate multiple replicates of multiple points across the entire claimed reportable range. We look at the line of regression and the different plots as recommended in Clinical Laboratory Standards Institute's guidelines on the evaluation of linearity of quantitative measurement procedures.

Now all IVDs have some interference from certain compounds, and blood glucose monitors are no exception. When we evaluate interference with glucose meters, we identify the substances that interfere and define the extent of the interference. There are common endogenous and exogenous substances that can interfere with glucose meters. Endogenous refers to substances naturally found in the patient's blood, such as cholesterol or bilirubin, and exogenous refers to substances such as drugs.

We evaluate endogenous and exogenous substances that have been known to interfere with glucose methodology. The endogenous substances, such as cholesterol, are evaluated at the highest levels at which they are known to occur in patients' blood. The exogenous substances, such as acetaminophen, are evaluated at therapeutic levels and at the highest levels at which toxic doses may occur. Samples are evaluated at clinically relative decision points, and the calculated bias, or difference, of the sample containing the interferent to the reference should not be greater than plus or minus 10 percent.

For hematocrit, we evaluate samples with glucose concentrations that span the measuring range at hematocrit levels that span the claimed range. We compare individual results of each hematocrit and glucose combination to results of a sample at the same glucose concentration with a normal hematocrit, and also to a reference value. The calculated bias should not be greater than plus or minus 15 percent for hematocrit interference.

Other factors that we evaluate are environmental effects such as temperature, humidity and altitude, and we also look at conformance to the International Electrical Commission Standards for medical electrical equipment. We also look at electromagnetic compatibility, and we look at software.

And lastly, FDA evaluates the labeling of blood glucose meters. The term labeling refers to all printed material that will be provided with the marketed device. This includes user manuals, test strip inserts, quality control solutions inserts, quick reference guides, if applicable, and also box

and container labels. As mentioned earlier, IVDs have their own labeling regulation, so when we evaluate glucose meter labeling, we make sure it contain the required elements, such as intended use, instruction for use and performance, as mentioned earlier. And since most glucose meters are for over-the-counter use, we evaluate the readability assessments of the labeling to ensure that it is written at no higher than eighth grade level.

So in summary, I have talked about many individual factors that can affect blood glucose meters. However, another factor that I haven't touched upon is user error. User error, intentionally or unintentionally, is misuse of the device, or when a user does not follow the instructions for use, or disregards the limitations and warnings in the labeling. This will be discussed in a separate session tomorrow. But the factors that I have discussed, and their effects on blood glucose monitoring systems, are those that can occur when the device is being used as intended, with the limitations heeded and the instructions appropriately followed.

And while we evaluate each of these factors individually, and determine the individual acceptable limits where appropriate, it is important to realized that the user experiences the cumulative effect of these factors. Thank you for your attention, and I think we have time for any questions.

DR. CLARKE: If you have any questions for Ms. Bernhardt, she is the FDA person who is speaking this morning, so please come forward and state your name and ask your question. (Pause) You had your chance. Okay.

The next speaker for this morning is Dr. Mitchell Scott, who is Professor of Pathology and Immunology at Washington University School of Medicine, and he is going to talk to us about analytical performance of blood glucose meters, and give us a state-of-the-art presentation. Dr. Scott.

DR. SCOTT: Well, first, I'd like to thank Courtney and Arlene for inviting me, and what I would like to do today is, first, figure out the computer.

I'll try to sort of give us a historical perspective on glucose meter evaluations. Now there's

no shortage of glucose meter evaluations in the literature, and if I stood here for 20 minutes, and went through 20 different glucose meter evaluation studies, I'm pretty sure everyone out here would be asleep by the end of it. So I'm just going to give a few representative examples, and then present some suggestions about, or actually go through the allowable error criteria that are there now. I'll talk about some data from my own institution, which is a 1,200 bed academic center, and then make some suggestions about maybe where we should start the discussion in terms of allowable error criteria.

So I think some of these things have already been stated this morning, the meters are getting smaller, there's over 30 of them listed on the ADA Website for home use, there's five manufacturers for hospital use. It's a growing market, and it actually accounts for, according to a Frost and Sullivan report, 30 percent of all laboratory revenue, in vitro diagnostics laboratory revenue. So it's significant.

Recent improvement in meters, no wipe strips, sample volume detection. Most importantly

from a hospital perspective is data storage and capture. It's only been five or six years that we've actually been able to capture the data from our glucose meters at our institution. Prior to that they were performed, they sometimes got written in the nursing notes, but we had no other way of capturing the data until five or six years ago. Alternate site testing was mentioned. There's a time lag issue in terms of alternate sites, and I'm not really going to discus that because we aren't using this in the hospital at this point in time.

Interferences, more this afternoon: the hematocrit effect, anemia, higher glucose. Keep that in mind when you think of these meters being used in a critical care setting. Anemia is very common in intensive care units. Reducing agents, the effects vary by method; I'll show you just a few figures. But some newer meters correct for hematocrit and reducing agents. And I think the theme that I'm going to try to present through this talk this morning is that meters are getting better, and I think they can beat current criteria that are out there.

So here's an example of the hematocrit effect, significant differences, particularly in hematocrits that you'll actually see in an intensive care setting, and if it's falsely elevated, is that going to cause overdosing of insulin? PO2 effect. Glucose-oxidase methods are affected, glucose dehydrogenase methods are not. Here's a glucose-oxidase method that takes it into account, and is not affected by PO2. So they're out there.

The maltose for the GDH methods has already been mentioned, as has user-induced errors. Now in this study, this was users actually performing self-monitoring following the package insert. And what they did in this study, it's somewhat contrived, I agree, but they had a mechanical flicker, that after the sample was applied to the strip, this little spring would hit the strip, and this is what happened to the results. So this was a sort of a reproducible way to induce user error. So you can see that just flicking the strip a little bit can greatly alter the values.

Okay. My main interest is the use of meters

in hospital setting, so I have to mention tight glycemic control. Today, most critical areas have protocols to keep glucose below now 130 until NICE sugar was generally below 110. This requires frequent and rapid glucose values, resulting in IV insulin adjustment.

Now what kind of an impact has TGC had on the use of glucose meters in hospitals? In 2000, before TGC became standard care, we used about 250,000 strips a year. Last year we used 550,000 strips at our institution, and almost half of those are in critical care settings. These meters are being used in intensive care units to the tune of 250,000 times a year at our institution alone.

So the original Van den Berghe study is blood gas instruments in arterial blood, the second use, the Hemo- Q, but very few of the TGC studies out there even tell you what method was used to measure glucose. The meta- analysis that was published in JAMA, in August of 2008, looked at 27 studies. Only ten of those studies, when I actually looked at the original papers, tell you how the glucose was

measured. Two of those were the Van den Berghe studies, that didn't use glucose meters, eight used meters, and seventeen, you have no clue, it does not say how glucose was measured. In the NICE-SUGAR paper you can't tell; it says, glucose was obtained by meters or blood gas analyzers, that's it. Again, I just want to point this again as the interference. Here's some meters are greatly affected with low hematocrits, and anemia is common in intensive care settings.

What about sample type? This is a study from Mayo, with 20 patients in the CCU, and they looked at the first five hourly samples after they began IV insulin. They looked at arterial, venous and capillary samples compared to the main lab. Now I'm not going to show the capillary data; it actually matched very well to the main laboratory. But arterial samples, the difference between the main lab and arterial blood gas values ranged from a positive 35 or so to a negative 5, but the average difference was a positive 15. Is that going to lead to over-usage of insulin if arterial samples are used?

How about venous samples? The same mean bias, but tremendously more scattered, using a venous sample. And again, the TGC studies that are out there do not usually say what type of sample was used. And I know in our critical care settings, it's all three, it's finger stick, its arterial and its venous.

So, after the JAMA meta-analysis we started thinking about this and wrote an editorial in Clin Chem, asking the question of whether or not these things are up to the task of use in critical care settings. Pointing out the fact that they are not to be used for diagnosis, so should they be used to make dosing decisions with what is a very dangerous drug?

Okay, accuracy and reproducibility. Where have we been, where are we now? We've seen these already, so I'm not going to spend a whole lot of time here. Main laboratory is 10 percent or 6, and these are the other criteria with the ADA, of course, being less than 5 percent, which all know really no meter can achieve. What I really want to point out here is the requirements of all of these criteria, that 95 percent of the values fulfill these criteria.

What about the other 5 percent? Okay, we do 600,000 a year at our institution, and let's just do a comparison study using one dear of data from my institution, okay. 600,000, that means 570,000 glucose values have to be within 20 percent. 30,000 can be anywhere. So I think we really need to think about this 5 percent of the values just sort of being unclassified.

Okay, let's just review a few representative studies, saying we'll start with some older ones. This was 2,000 values performed by nurses, published in 2001 but it was really data from the 90s. And if you look here, you'll see that 25, 15 percent of the values are not within 15 percent. So pretty poor performance in the 90s.

This is a more recent CDC study performed by Mary Kimberly and Gary Myers here in the front row, at CDC. A single medical technologist operator looked at five meters to see ranged from 6 to 11 percent, even more important, up to 32 percent bias between meters. And their conclusion was that standardization was necessary for meters and they needed to match the main

laboratory much better than they currently did.

This study came from the NIH, was commissioned by the FDA of four common meters, and the data was a little bit better, CVs in the one to nine percent range; again, health care operators, not self-users. The worse CVs, as expected, were at lower glucose values, but still, eight or nine percent, that's not too bad below 60. But again, there were significant biases to the laboratory reference method, so between negative 12 and positive 12 per cent bias, so up to 24 percent.

Now let's take that study and look at the various criteria. Okay. All four of these meters would not have done very well with the ADA criteria, 92 percent of the values would have failed with meter A, 42 percent with meter C. Meter B, C and D all would have done quite well with the current FDA and ISO criteria, and very well with the Clarke error grid, including meter A, which was miserable by any of the criteria up above.

So, which criteria should we be using? I don't have an answer; I'm going to end with a

suggestion. Let's look at CAP data. There is where we are today. See these between methods, and this is all hospital meters with five manufacturers represented here. With exception of a few, where the end is very small, the CVs with very large ends is not that bad any more. I mean, it's not main laboratory CV. If you look at this same report from the main laboratory it's between 1 and 2 percent, but a lot of 3 percent CVs between all laboratories using the same method.

Now for CAP data it's a little unfair, I think, to compare mean values per system, because of matrix effects. So I think the 78 to 124 difference between meters is probably partially due to matrix effects. But let's look at the low and high value within a system. Here's a meter used at 4,000 institutions, and the lowest value was 83, the highest was 116. That's within a system, so there is still spread within a system. But overall, I think the CVs are reasonable.

Okay, what about Barnes Jewish Hospital at Washington University Medical Center? We've got about

200 meters out on the floors, over 2,000 certified operators. We do 1,400 of these strips a day, and this is our QC data, very similar to the CAP. I think, you know, 200 meters, 4,000 QC points a month, CVs in the 9 to 6 percent range. That's how we do in our institution. I actually distributed a sample to 40 different sites, same patient sample, and the range of values was 153 to 210 in a 95 percent confidence interval of 139 to 195. So there is still some, with the system that we're current using at Barnes, considerable spread.

Now this data from our institution I think is particularly telling. I get a report on a monthly basis of all meter values that are repeated within 15 minutes. The reason for this report is actually a billing compliance issue. We can't bill for duplicate testing. So I saw it, not from a compliance billing opportunity, but let's see what kind of repeat values we're actually seeing. And this number is between .8 and 1.2 percent of all of our glucose values every month. This just happens to be November, 2009. So we did about 40,000 or so that month. 357 were repeated

within 15 minutes. The median repeat time is 3.5 minutes.

So these are not values being repeated because they just gave insulin or they just gave orange juice and they want to see what happens. These are values being repeated because the operator didn't believe the number. The mean absolute difference between repeat values is 84, in November. That number I've seen as high as 125. These are clearly being repeated because they didn't believe the first value. What about values that are performed and never repeated, because they're close enough. I have no idea of that number, but I think there's a hint in these repeat values here.

Okay, let's go back to some comparison studies and look at some more recent ones, published last year. This is a newer meter. All the CVs in this evaluation using venous arterial whole blood and aqueous Qz material were less than 5 percent. No effects from hematocrits or reducing substances, this is one of the newer meters that corrects for that. Very good bias values.

If you put it on an error grid, that is what that particular study looked like, fairly impressive. This tells me that probably some of the newer meters are able to pull off tighter performance criteria than what's currently out there.

Here's another study of a new meter. Again, very good precision and very small bias, you see, plus or minus 10 or so versus the laboratory method.

Okay, so what criteria should we use? I think the outcome studies to determine that are probably not going to happen. It's going to involve randomizing patients to a meter versus a main laboratory or a blood gas analyzer, and the end's going to have to be tremendously large to detect a difference, I believe.

But, how about biologic variability? Callum Fraser is a very strong proponent of using biologic variability. And for glucose, total allowable error should be less than 8 percent. Well, now, what's the biologic variability in a critical care setting? I have no idea, I don't think anybody does. But, is this a place to start?

So, what I'd like to end with is this is a discussion point and a place to start. 10 percent, or 10 milligrams, I think the newer meters can probably do that. And then let's develop a new error grid for critical care settings, and George Cembrowski and Jon Krouwer are big proponents of doing this, and I think you can. I don't have the crayon skills of Dr. Clarke, so I didn't actually draw a new error grid. So 95 percent within 10 percent, or 10, would be like zone A, 99 percent of the values within 15 percent, or 12, zone B, and then nothing exceeding 20 or 15, or very few. But I think what this does it accounts for all the values, and you don't have this 5 percent sort of random numbers out there that can be anywhere, according to the current criteria.

So, what about -- can new meters do this? Well, I don't know. Here's another study in 2009. This is the current ISO criteria, and I just drew in here the 10 percent, maybe not. But I think it's a point to start discussions and see where we end up.

So with that, I believe the newer meters appear to be getting better, the comparison studies

seem to suggest better in precision, less bias, less effects from interferences. I think they can clearly do better than the current ISO/CLSI criteria.

Can they meet a 10 to 15 percent allowable error? I think time will tell, and the FDA presentation before me showed that half of them, at least, probably can. But, in the overall analytical performance criteria, you've got to always consider more than just precision and bias. You've got think about interferences, particularly in critical care settings, sample types, user errors. These all go into the evaluation of a meter and have to be considered.

I believe the patients receiving a dangerous drug should have the best analytic method available, and, quite frankly, there are alternatives to meters in critical care settings, blood gas analyzers, a few meters that are approved for diagnosis. They probably cost more, but they're out there. So with that I'll be glad to take questions.

DR. CLARKE: Questions?

DR. GINSBERG: Dr. Barry Ginsberg, Wyckoff,

N.J. You sought to raise an interesting question, which I'm going to raise in my talk as well, which is that of outliers. And it's interesting, because when you look at the clinical data from virtually every meter out there, which is done on 100 or 200 or 500 patients, you don't see outliers. When they give their ranges on Clarke or the Consensus error grid, its 97, 98, 99 percent A, and 2, 3 or 4 percent B, no Cs, no Ds, no Es.

When you look at data from the manufacturers when they're not doing 100, 200 or 300, but in the laboratory they're doing 5,000, 10,000, 15,000, 50,000 numbers, you start to see that outliers are real. I'll show you some data on that. That depends upon the company, but somewhere between .03 and .1 percent of data is out in that range, is which is more than 100 or 200 milligrams away. And it's real, and it's something that I think we need to start thinking about.

DR. SCOTT: Yes, I mean that's exactly why I presented the duplicate data from our own institution, where almost 1 percent are being repeated, and of

those, probably right, .2, .3 percent show greater than 150, 200, I mean, they go as high as a 450 difference between values.

DR. ELAHI: Dariush Elahi, Baltimore. Could you make a comment? When you spoke about repeatability, you were speaking of a second sample within 15 minutes, I assume.

DR. SCOTT: Correct.

DR. ELAHI: Could make a comment about the repeatability of this exact same sample within a minute of duplicate reading or triplicate reading. What are your thoughts about that? How close should that be?

I took a blood sample, I don't believe it. Maybe I made a mistake, I'm going to do it again. Now I have a 40 milligrams difference. Should I take a new sample? Should I run the third time?

DR. SCOTT: The problem is these decisions are being made by the patient caregivers.

DR. ELAHI: I'm asking your opinion.

DR. SCOTT: Oh, my opinion? I think if it was carefully performed, they're probably going to

come within 5 or 6 percent of each other, which is what our quality control, what the studies show as current imprecision of glucose meters. What we're seeing with our values that are repeated, in my opinion are user errors, not an error by the device itself. But something was done incorrectly by the user.

DR. GINSBERG: I disagree with that. Again, using some data, in 1993 we tested about 1,000 strips, having the patients test their own meter against a reference. And we found that the average error was about 13 percent. And it turned out that at that time, if a health care professional did it, they had an average error of about 7 percent. So in the early 90s, strips, required a lot of attention to how the user did it, and professionals were much better than patients.

We repeated that study in 1996, and what we found was that the professionals were still about 7 percent, a little bit better, meters and strips had gotten a little bit better but not a lot, in the hands of people who were very careful about how they did it.

The difference was at that point, that patients were substantially more accurate than professionals, that patient's were about a half percent better than the professionals were at that, because they actually do it more often. And I think, the outliers I talked about were done in the laboratory by highly experienced professionals. I think there are some, whenever you produce large numbers of anything you're going to have some outliers.

DR. SOLDO: Monnett Soldo. I'm from California. My question is, it seems that there are two issues here, one is whether or not meters can meet the accuracy of the ISO or any other standard in any environment, and the other is in an ICU environment.

And for one, with all the data that is being presented, I personally am having trouble distinguishing between these two. I can't tell when you present data, and when any of the panelists present data, whether it's for an ICU setting with significant interferences, or whether it's just what the manufacturer did. So, to the extent possible, if we can distinguish between what data is being

presented that's relevant to an ICU setting, or point-of-care setting, I think that would be helpful.

DR. SCOTT: The only two studies, one I said the Mayo study was clearly from a CCU, the evaluation of the newer meter strip in 2009, that was also from a critical care setting and I failed to say that. The others were performed in the laboratory or by users.

MS. MANN: Hi, Elizabeth Mann. I'm from the Army Burn Center. We tested five different meters, four of the currently available meters, and then the new 4- channel meter. I personally did all the glucose measurements myself. I'm a nurse, throughout our ICUs within our hospital, Burn, Trauma, Surgical and Medical ICUs.

The error in normal patients who have hematocrits between 18 and 25 percent, generally, after the ABARE study, where we definitely titrate to a very low level, I think this point about was it done in the ICU versus non- ICU is extremely relevant, especially with the data that you presented here.

I would just like to say that we did develop a correction factor for that hematocrit. It's a

simple math formula and we applied it to each of those meters, and we were able to achieve less than 5 percent error in real patients. And then we did the 4-channel glucometer and it was no different in our labs. So I would agree with you that, yes, the technology is there with the new meter, but even more so, the technology is actually available in the current meters if they're corrected. And when you correct, you actually capture occult hypoglycemia, and we found that we reduced our hypoglycemia rate within our unit by just treating the proper number with insulin.

DR. SCOTT: I believe there to be some practical issues, unless you could actually get the meter to do it.

MS. MANN: Well, that's the best thing. Right now we use the computer and the nurses enter the glucometer value into a computer, so it is an extra step, but we felt like, for safety's sake, when these glucometers all had errors more than 15 percent, some over 20 percent, it was simply not safe to do tight glycemic control without some sort of correction until

we got the 4-channel glucometer.

DR. SCOTT: We've got over 2,000 operators, and trying to get sort of a manual step conversion like that, I think it would be very difficult, really.

MS. MANN: As a nurse, just to follow along with the user error, I do agree. When we're busy, and we're doing tight glycemic control in this day and age, and we're doing a measurement every hour, the strips sit for a long time, there's not appropriate draw out of the A-line or the central line. So I agree, it really is user error rather than repeatability on a device.

DR. CLARKE: Thank you, and we will proceed. Our next speaker this morning also comes to us by way of Washington University. Dr. David Sacks is Associate Professor of Pathology at Harvard, and Direction of the Clinical Chemistry Division at Brigham and Women's Hospital. He did his medical training in Cape Town, South Africa, and Residency in internal medicine at Georgetown, and Clinical Pathology at Washington University School of Medicine. He is a Fellow of the Royal College of Pathologists,

and is currently chair of the National Glycohemoglobin Standardization Program Steering Committee. And he is going to talk to us this morning about the clinical perspective, and the clinical need for tighter performance requirements. Dr. Sacks.

DR. SACKS: Thank you for the introduction, and I'd also like to thank Courtney and Arlene for inviting me to speak here today. So if I can get this thing to work. So just an overview of my talk. I'm going to talk about, since I'm doing the clinical perspective, I'm going to give a little background on the clinical aspects, the need for near normal glycemia, then I'm going to divide the clinical use of meters into two groups, self- monitoring of blood glucose and tight glycemic control in ICUs.

And so I thought, being the first clinical speaker, I should give a little bit of context. Now this has been touched on before but this is a map of the world, showing the prevalence of diabetes in the year 2000, which are the red bars, and the blue bars are the predicted prevalence in 2030. And you can see, in the United States there's a predicted increase

of 70 percent. But the most dramatic increases are in Asia, India and South America, where the increases are up to 150 percent or more.

And, of course, the major complications of diabetes, which is the reason why people with diabetes are treated, are microvascular complications, retinopathy, neuropathy, nephropathy, and then, of course, the macrovascular complications. So I'd just like to emphasize the Diabetes Control and Complications Trial, because that was the major study that looked at the role of self- monitoring in diabetes, and they had 1,441 patients, all with Type I diabetes. And they were randomized to either intensive or conventional insulin therapy. And the goals of the intensive therapy were evaluated by the patients performing self-monitoring four times a day, and also monthly hemoglobin A1Cs. And the patients were followed for a mean of 6.5 years.

And the outcomes of the study were very dramatic. The patients on conventional control, this is the glucose, capillary glucose, during a single day at different times of the day. The conventional

control is the top line, the intensive control is the lower line, and you can see there was a marked reduction in glucose values at all times of the day in the intensively controlled patients.

And if one looks at the development of complications, the conventionally controlled group are the blue bars, the intensive control the red bars, all the microvascular complications, retinopathy, nephropathy and neuropathy, were significantly lower in patients on the intensive control. Macrovascular complications were not decreased at the time the study ended in 1993, but subsequent follow-up has shown a 50 percent reduction in myocardial infarction in the intensively treated group.

So what are the recommendations for portable meter use? Well, they're used by patients at home, work and school, and also in acute and chronic care facilities, including ICUs and in physician offices.

So I am going to talk first about self-monitoring of blood glucose. And like Dr. Scott, I'm going to pick a few representative studies again, just to illustrate points. So self-monitoring is

recommended for the following reasons:

To determine the insulin doses at meals, and the recommendation by the ADA is this should be performed at least three times daily if people are on multiple insulin injections. And there is evidence in the literature that glycemic control is worse if this is done less than three times a day.

To determine the insulin dose in pregnant women with diabetes.

To determine whether patients are achieving glycemic controls.

And the last one, which is really important, is detection and avoidance of hypoglycemia, and actually there is a very large body of literature on the fear of hypoglycemia. This is the factor that's most limiting for tight glycemic control.

So that raises the question of does meter performance meet clinical needs? So let's have a look at this. So hypoglycemia, the risk increases dramatically with therapy directed at near normoglycemia. Patients in the intensively treated group in the DCCT had three times the rate of severe

hypoglycemia. As I mentioned, near normoglycemia reduces the risk of chronic complications, so this is the goal.

Another important factor to consider is that many patients with Type I, and even with longstanding Type II diabetes, have hypoglycemia unawareness, so they lose the symptoms of hypoglycemia. And self-monitoring is the only defense when the symptoms are lost, because the patients don't know they're hypoglycemic. And severe hypoglycemia is associated with mortality, dementia and harm to self, or others, for example, if the subject has an episode while driving.

So there are a lot of books about hypoglycemia. If one looks at the internet you can find the recipe to conquering hypoglycemia. You can even find hypoglycemia for dummies. But the important question is, can meters reliably detect hypoglycemia, since this is one of the major uses for people with diabetes. Say, if one uses the current ISO/CLSI criteria, and you imagine that an individual has a true glucose of 50, if one does the calculation the

acceptable limits, the results, are 35 to 65 milligrams per deciliter. So 35 would be severe hypoglycemia, 65 would be, you don't have to do anything.

And of courses, as was emphasized in the previous talk, 5 percent of values may be outside of these ranges. The problem is, the patient may not know which values are wrong. So if the patient ends up with a value that's outside the limits, they may not know, particularly if there is hypoglycemia unawareness. So I would argue that these results cannot reliably detect hypoglycemia. And you may end up with severe hypoglycemia, requiring help from another individual, and again, somebody may actually have to call 911.

Let's look at meter use in practice. Accuracy criteria are exclusively for analytical performance, as has been mentioned before.

I think one of the advantages of being the third or fourth speaker is that there is limited overlap with what's gone before. The last speaker, I don't know if they're going to be able to say

anything, because probably everything they need to say will have been said before. So I apologize for some repetition. But it helps, you know. People doze off, you know, you wake up. For those of you who stay awake all the time, I'm sorry.

But in practice, the accuracy criteria did not consider pre- or post-analytical error, and very importantly, the evaluation is usually performed by highly trained medical technologists under optimal conditions. And the ISO specifications, I would argue, do not inform the condition of how the meters actually perform in the patients' hands.

So as it's been mentioned, current meters have performance superior to prior generations, you know, the technological advances by the manufacturers have been remarkable, they decrease operator error. But the evidence in the literature shows that performance by patients is inferior to medical technologists.

I'm just going to give one representative example of this study from Norway, by Sandberg. And they looked at five different meter types, and the CVs

for patients among these five meters ranged from 7 to as high as 20 percent, and the medical technologists had CVs of 2.5 to 5.9 percent. So there's a huge difference in the hands of patients who were performing self-monitoring anyway. These were people who were doing self-monitoring.

And this is an example of one of those meters, this is from the same study, what's shown in the vertical axis is the percent deviation from the laboratory method, the horizontal axis is the concentration of glucose in the lab method. The left panel is patients, the right is technologists, and you can see a huge difference. This meter seems to have a positive bias, even in the hands of the medical technologists, but there's a very tight precision when it's done by technologist.

And one of the interesting findings in the study, which was published in 2002, is that patients failed to meet ISO criteria. None of the five meters in the patients hands met ISO criteria in this study. Two of them met ISO criteria in the technologists' hands, but none in the patients' hands.

How accurate does a glucose measurement need to be? Well, this has been addressed a little bit before. Several criteria have been proposed for analytical goals. There is expert opinion, there's opinion of clinicians -- I don't know what the difference is between 1 and 2 -- it implies that clinicians are not experts, state-of-the- art, regulation, biological variation. But there is no consensus regarding this.

But what's missing from this is what do patients think? So for most, analytes were measured in the lab, it's done in the clinical lab. But self-monitoring of glucose is different because this is actually done by patients themselves. So what do patients think? How accurate do patients think that the meters need to be?

So there was a very nice study which looked at 201 patients with Type I diabetes. And these people, on average, had been doing self-monitoring for ten years, and there was a means of 11 measurements a week, which is less than the recommendation, but still, people who had been doing this for a long time.

And each individual completed a questionnaire, there were nine questions. And the patients said that they reacted to critical differences, which is the difference between two consecutive results, of 22 to 30 percent. So that's quite a big difference with four patients reacted.

And this shows you, from that study, the vertical axis is the imprecision, the horizontal axis is the bias, and the critical difference of 22 percent is plotted here, and 30 percent is plotted there. So based on the results of the study, the calculations, the analytical CV of 6.4 to 9.7 percent is needed for patient-derived criteria. However, this excludes hypoglycemia. And the patients were very afraid of hypoglycemia. And for hypoglycemia, they wanted a CV of 3.1 percent, which is incredibly low, but this is what patients say they need. And based on the study, a CV of less than 5 percent and a bias of less than 5 percent are required to meet the expectations of 75 percent of these patients. Again, this excludes hypoglycemia. So patients expect their meters to perform very well.

Now in 1987, the ADA recommended a total error of less than 10 percent for 100 percent of results, and in 1996, in response to the DCCT, the ADA modified this to total analytical error of less than 5 percent. And as has been mentioned, the other recommendations, CLSI or ISO, 95 percent of results, and this has been discussed in some detail, but I'm going to discuss it again. So, as was suggested by Dr. Scott, I think we need an addendum to meet a performance criteria, because 5 percent are excluded from accuracy criteria. And these values can be essentially anything. So, if you do the calculation, if a patient does self-monitoring of blood glucose four times a day, you'd expect one result to be outside the analytical limit every five days. The problem is the patient won't know which this one result is, which is outside. So that's very, very frequent. So I think we need to define criteria that include these 5 percent of values.

Okay, so now I'm going to talk about tight glycemic control. I like this picture, which was put in a News and Views article I wrote by the editors of

Nature.

So does meter performance meet the clinical needs in tight glycemic control? Nobody can see this of course, you're not supposed to. This is from that JAMA meta- analysis. So there're lots and lots and lots of studies, so how does one pick a study to evaluate this? And I thought I would pick the NICE-SUGAR because that's the study that seems to have influenced practice the most since the original paper by Greet Van den Berghe.

See, if you're from South Africa you can pronounce the name properly. I think it's close; there's somebody here from Belgium that could correct me.

But it was also the biggest study. And for those of you who don't know, it's a multinational study designed to test the hypothesis that intensive glucose control reduces mortality in 90 days, a very large study, 6,104 adults, who were admitted either to the medical or the surgical ICU at one of 42 hospitals in Australia, New Zealand or Canada. And within 24 hours of admission, the patients who were expected to

require more than three days, at least three days of critical care, were randomized either to intensive or conventional glucose control. And they had quite tight criteria. They target glucose ranges where intensive control was 81 to 108, and the conventional control was less than 180 milligrams per deciliter.

So what were the results of the study? Well, if you compare the intensive and the conventional treatments, there was a significantly lower mean glucose in the intensively treated group, 115, versus 144 in the conventional group. There was a statistically significant increase in mortality in the intensively treated patients, 27.5 percent, versus 24.9 in the conventional group. And a remarkable increase in severe hypoglycemia, which is defined as less than or equal to 40 milligrams per deciliter, 6.8 percent in intensive, and 0.5 percent in the conventional group.

Now this is hypoglycemia that was detected. So as Mitch mentioned earlier, it's impossible to obtain details concerning how glucose was analyzed in this NICE- SUGAR study. So this has changed clinical

practice, but we don't even know how they measured glucose.

What did they say? They said glucose measurements were performed on arterial blood whenever possible -- I don't know what that means -- using point-of-care or arterial blood gas analyzers, or laboratory analyzers in routine use at each center. I mean, this is just absolutely useless information. There's just no way to know how glucose was measured. And this has changed clinical practice.

So let's look at the study in a little more detail. So, one of the potential problems is that the different glucose values produced by diverse methods and samples will lead to different insulin doses, and potentially wide variations in true glucose concentrations, right, because the results they're getting may not be the true glucose concentration, which is what we really need to know.

So I did some analysis of this study. Now if you look at CAP proficiency testing dotted for 17 meters, the bias is up to 41 percent. So let's think about that for a little bit. So if the true glucose

is 144, which is the mean of the conventionally treated patients in the NICE-SUGAR study, a bias of 41 percent is 59 milligrams per deciliter. So if you take that value, the difference in the mean glucose between intensive and conventional groups in NICE-SUGAR was 29 milligrams per deciliter, so the bias can be twice the difference in the mean glucose between these two groups. That's just bias.

So if you think about this, if a meter is high bias, what does it mean? The results will be higher than the patient's actual glucose, and the patient will receive too much insulin and will develop hypoglycemia, which may not be detected in these patients because many of them are unconscious. And we have no way of knowing whether these people are hypoglycemic if you use a meter that has high bias.

Let's do a little more analysis of the study. So let's say the patient has a true glucose of 95, that's right in the middle of the target goal for the intensively treated patients. So the acceptable range for the meter, based on current ISO/CLSI criteria, is 76 to 114 milligrams per deciliter. So

the meter will give you this result. Anywhere is acceptable between 76 and 114. This is excluding any of the conditions that occur in patients in ICUs. Remember, 5 percent of values could be outside this range, and when you're doing hundreds of thousands of these, a lot are going to be outside the range. Now these values, 76 to 114, exceed the range for the intensive control target of 81 to 108 milligrams per deciliter. So I'm not sure that anybody in this study knew what the actual glucose was.

So, I'm going to summarize now. Measurement of blood glucose concentrations has a very important role in patient management, obviously. Accurate identification of hyperglycemia is absolutely essentially for people with diabetes who are doing self-monitoring, and certainly patients in intensive care units on tight glycemic control protocols do. Current performance criteria for glucose meters, I would argue, are inadequate for clinical needs. And based on biological variability, the recommended accuracy criteria for meters would be a minimum of plus or minus 15 percent, desirable would be plus or

minus 10 percent, and optimum would be plus or minus 5 percent, clearly not achievable with the current generation of meters.

And I'd just like to end with a potential new problems with meters that has come up recently, that people would not have thought of before, and this is somebody on an airplane that goes beep-beep-beep and he says, sorry, it's just my glucose monitor. Thank you very much.

DR. CLARKE: Go ahead, please.

MR. ERVIN: Ken Ervin from California. Dr. Sacks, given the known differences in the way CAP proficiency samples behave with respect to glucose meters, do you think that's an appropriate way to evaluate bias of glucose meters with old whole samples?

DR. SACKS: Look, clearly, I mean one of the problems with a CAP proficiency testing for glucose meters is that it's not whole blood. I'm actually on the CAP Committee, Chemistry Committee, and every few years I raise the point of can we do whole blood, and we have not been able to do that. So clearly, some of

the bias is exaggerated in that study, as you point out, because of matrix effects among the different meters, which was mentioned earlier by Dr. Scott.

DR. CLARKE: Other questions? Well, I have a question. Would you suggest that our criteria for testing meters and looking at their accuracy should include a specific percentage of numbers in the hypoglycemic range? In other words, we see lots of data presented on new meters, but very few of the data points are less than 70 milligrams per deciliter, whereas when we're out in the real world, what we see is 8, 9, 10 percent of our patients' glucose numbers are actually in that range.

DR. SACKS: Right. Now I think that's a very good point that you make, and I think that since the hypoglycemic range is so critical to patient decision making, I think it's very important that the meters perform as accurately as possible in that range, and they should be evaluated for that.

DR. CLARKE: Go ahead, please.

DR. SIMMONS: Dr. David Simmons, Tarrytown, NY. Professor Sacks, we had this conversation at the

ISO meeting in February, and I'm still struggling a little bit with how we can take a study where you led, NICE-SUGAR, led by telling us we have no idea what samples were used or what analytic method was used, take the results of that study and then very elegant analysis of what the potential modeling may have been, but use that as a strong argument to say that hand-held meters, which might be used for home care, or might be used in the study, but we don't know whether they were used in this study, use that as a compelling story for why hand- held meters might not have the analytic strength to be used in intensive care units.

So my concerns are, one, we haven't separated out home use from intensive care unit use; and two, we have no idea whether the study cited, with the problematic outcome, had any use of the meters where you used the modeling from a 41 percent bias, which I don't think that the current standards would accept, would provide inadequate results. So I think that I'm seeing a lot of mixing of different results to come to the conclusions.

DR. SACKS: I'm sorry you were confused. I

thought I made it very clear with the difference between -- the first half of my talk was on self-monitoring, and the second half was on tight glycemic control.

And what I was trying to point out in the analysis of the NICE-SUGAR study is, as you indicated, clearly not only glucose meters that were used in the study. But I'm just pointing out the limitations of the glucose meters as currently accepted for ISO criteria could not be used properly in a study of that nature because of the large variation in acceptable criteria for meter accuracy.

DR. SIMMONS: But I just think, to be completely transparent, we should de-link the results of NICE-SUGAR from those criteria, because we don't know what analytic methods were used. And then the other is there are other issues with NICE-SUGAR, including attribution of morbidity and mortality in the time windows applied. So although the analysis may be very interesting, I think we need to de-link that from the outcome in NICE-SUGAR, to be completely transparent.

DR. SACKS: I'm not sure that some of the, that you can actually do that, because it's quite possible that some of the patients in NICE-SUGAR, some of the worst outcomes, may be due to undetected hypoglycemia. And there's no way to know whether that's true or not.

MR. COMBS: Art Combs, St. Louis. I'm actually an ICU director. ICU patients have nothing to do with ambulatory patients. I was interested in this a number of years ago and I just looked at the laboratory printout and found that more than 50 percent of all of my admissions had at least one abnormal ionized calcium. More than 80 percent had at least one abnormality serum phosphorous. These are things you never see in an ambulatory person. So that's one point I wanted to make.

The second I wanted to make is that we heard at the beginning, from the regulatory perspective, that this is all about intended use. I would ask the panel what in home care is intended for the intensive care unit, and the answer is nothing. And the standards for the way we measure blood pressure in

home care, or the way we measure temperature in home care, or the way we measure anything in home care, has nothing to do with the intensive care unit.

I would submit to you that in the intensive care unit, there currently are no adequate technologies to test the hypothesis of the benefits of tight glycemic control. With that conclusion, I agree 100 percent. But to suggest that every ambulatory diabetic should be walking around with an instrument qualified for the care of the critically ill is preposterous. I don't think we need standards for home care that were established in the intensive care unit. We need to divide these, we need to talk about them completely separately, and we need to understand, in my opinion, ironically, the FDA's point of view, which is what is the intended us. Thank you.

DR. SACKS: Those are clearly very valid points, and I'm not for a femtosecond suggesting that everybody who dies in these ICU and tight glycemic control protocols is due to hypoglycemia. These people are really, really sick, as was made evident. So one of the suggestions that comes up from -- and

the other point, of course, is that everybody in this room knows is that glucose meters are not FDA approved for use in tight glycemic control protocols in ICUs. But having said that, people are using them, and they are used very, very, very widely, and not just in this country.

I mean, so one of the possibilities from your argument is that there can be two different criteria, one for meters in ICUs, and one for meters in home use. And that's something that is a topic for extensive discussion, and could certainly be addressed by the FDA and other people.

DR. CLARKE: Final question.

MR. COMBS: May I just offer a minor rebuttal now?

DR. CLARKE: What's a -- I'm not sure what a minor rebuttal is. The term, minor ...

MR. COMBS: Let me call it a clarification. It's my opinion, and this is an opinion, that the migration of hand-held glucose meters into the intensive care unit was, as many things in medicine, the result of a lot of perverse incentives. Some of

them actually, however, are legitimate. The turnaround time of laboratories may not be enough to meet the needs of the care of critically ill patients. In some cases there is point-of-care meters testing, in some cases there are satellite laboratories, in other places there really is only a central laboratory, and the turnaround time is simply unacceptable. And so the nurses wanted to have something at the bedside.

Secondly, there was the convenience and the cost factor. It is easier to put a meter at each bedside than it is to install, qualify, staff, et cetera, a satellite laboratory, and so that is another reason.

And the third is a very important thing, and I want to echo the comments of the nurse who spoke earlier. No one has any idea what a burden this is. Let's say, soup to nuts, from obtaining a blood sample, purging the arterial line, getting the strip, measuring the number, writing it down, re-purging the lines, establishing everything, takes five minutes. My nurses work 12 hours. Twelve times five is 60. That is

one hour out of an entire nursing shift, one hour, devoted to a single task.

I would submit that intensive care patients are a good deal more complex than that. So I believe that cost, convenience, turnaround time and immediacy or immediate availability of results are the things that drove hand-held meters to the bedside. But they don't belong there. They belong in the hands of ambulatory diabetics who are not on a continuous insulin drip, and who are not a priori critically ill.

DR. SACKS: So I have no rebuttal to your rebuttal because I agree with you.

DR. CLARKE: We are going to take a break until eleven o'clock, and you can enjoy your coffee and whatever in the foyer. Please take your seats so that we can keep on time. (Meeting recessed.)

SECOND SESSION

DR. CLARKE: The first speaker for the second session here is here -- ah, yes, oh, you were hiding behind there, Marc -- Dr. Marc Breton, who is in -- at the University of Virginia, has a PhD in Systems Engineering, and has specialized in the

application of engineering to medicine and has done human testing of autonomic or -- automatic, excuse me, insulin control systems and simulation of glucose insulin dynamics in man. And he's going to talk to us about clinical perspectives, self-monitoring of blood glucose inaccuracy, and clinical consequences in Type I using an in silico model. Dr. Breton?

DR. BRETON: Good morning. So I'm put in the slightly odd position of being an engineer addressing a mainly clinical panel and talking about clinical consequences, but I'm going to try to make as much sense as I can.

So, the quick background, which has been talked about at length by now, is most -- all Type I diabetics, and consequent number of Type II diabetics, who use insulin, face a challenge of trying to attain tight glucose control and avoid hypoglycemia at the same time. And they are trying to attain that tight glucose control for the reasons that were discussed previously and the consequences of not having a tight glucose control.

And the issue with working towards that

tight glucose control, that target, as low as possible is that you need a very good information, a very accurate information, about the status of the system to control it. And in that particular case, you need very good information about the glucose level of a particular patient.

Now, the only means for diabetic patients to have regular, frequent information about their glucose levels is the glucose meters. Now, Dr. Clark, on my right, in the review article stated the following sentence, which basically say that accuracy should only be seen in terms of clinical consequences, and that pure engineering study of the devices, though interesting, were mostly irrelevant to the patients themselves, and I tend to agree with that particular statement.

There were several studies along the last ten years, we'll say, that looked at particular errors of glucose meters and their clinical consequences. The first one that I wanted to cite was by Raine, et al. in 2008, and they studied the effect of miscoding meters. And they had a series of automatically coding

meters and non- automatically that were either coded or miscoded. And he looked at the consequences in terms of glucose excursions and insulin dosing and detected that the risk for hypoglycemia in miscoding meters, so creating a bias, was about ten percent higher.

There were very few clinical studies of the effect of accuracy, mainly because to design such a study is extremely intricate, and in some cases unethical. It's difficult to send a patient out in the field and then intentionally miscode or bias his or her meter by a hundred points.

So to circumvent that particular hurdle, what Burnt, Burns and Boyd did a few years ago, and what -- in a landmark article for us was that they used simulation, or a mathematical model of a human -- of the glucose response of a human being to study more thoroughly the effect of inaccuracies in meters. And the work that I'm going to present here is really in line with Dr. Bruns' work.

Now, the first thing I need to talk to you about, and I will try to be as short as I can, is

simulation, and the tool that I basically use in the study of this image inaccuracy. So what you're seeing here has absolutely nothing to do with diabetes, obviously. And the one thing that is interesting about these three pictures -- it was more interesting than last year -- until last year, the plane that you're seeing on the screen had never flew. And the Boeing 787 was entirely designed based on simulations, mathematical modeling, and such engineering tools.

So I present that slide to basically show you that simulation is not, by any stretch of the imagination, a new idea. It's a tool that has been used over the last decades by engineers to design and to test the validity of different systems.

Now, why do you need to create such a simulation in man and to apply it to diabetes? Well, first you need mathematical models that are based on clinical knowledge. And so some understanding of the physiology that will allow you to create mathematical models. You need to accumulate data targeting a specific subsystem of that physiology to understand how the different fluxes interact. You need to

identify the physiological processes and the fluxes that I just talked about. Finally, you create in silico subjects, and based on these in silico subjects, in silico standing in for a computer, you assess the intra-subject variability and create in silico populations.

Finally, you implement it in a software, and you absolutely need to validate it against in vivo data, and in vivo data that you did not use to create the simulation environment in the first place. And that, basically, allows you to have a software that will give you the opportunity to run what we call in silico experiments. The advantage of in silico experiments versus in vivo experiments is not only do you go incredibly faster, a fraction of a second to simulate several days of diabetes treatment, but you can test situations that would put the human at risk, knowing that in a computer, the risk is also mitigated.

Now, what you start with is that mathematical models based on understanding of the physiology. That work was started probably 30 years

ago, if not 40 years ago, by Dr. Cobelli in Padova in Italy, and his assistant, Kadava(ph) Amann. They had two publications in 2006 and 2007 presenting the mathematical models that we worked on. And then Dr. Kovatchev(ph) and myself took these models one extra step, and developed them to specifically address Type I diabetes. And you can see that you basically had mathematical models of the liver, the glucose system, muscle adipose tissues, insulin delivery. In Type II or normal subjects, you would have insulin secretion also. And, of course, the gastrointestinal tract to simulate meals.

You attach simulate measurements to these mathematical models. These allow you to measure the glucose of these patients. You can, of course, simulate the YSR(ph) of Beckman, or you can simulate an SMBG, which is exactly what we're going to do today.

We can also simulate a continuous monitor, which has allowed us to do closed-loop trials in simulation environments. You devise a treatment, obviously, and then you administer the treatment

through a simulated pump or a simulated multiple daily injection or even IV injection in the types of -- in the case of simulating ICU studies.

You need to accumulate data, and we accumulate -- well, we got access to 350 healthy Type I, Type II, and pre-diabetic patients across numerous universities and studies. For all these patients, they had one characteristic in common: all these experiments were done with triple tracers, and this gave us access to the fluxes of glucose and not only its concentration, so we actually know what's coming out of the liver, what's coming from the meal, which is used in the tissues, what's excreted. And using all these fluxes, we can actually identify a patient, and it corresponds to our good-old mathematical problem of the holes in the bathtub.

Now, the bathtub in this case is that particularly ugly depiction of -- you can, for example, recognize here the concentration of glucose in plasma. These are the insulin states. Insulin is administered subcutaneously in that case, makes its way to the plasma. Plasma and insulin influence the

different musage(ph) of glucose.

And what's even more ugly is the mathematical equations that correspond to that figures. I'm showing that first to justify my PhD, and second to show you that a patient is -- corresponds to 26 independent barometers. And I'm not trying to throw 26 to impress you. What I'm trying to convey is that it's absolutely impossible, even knowing perfectly the patient, to guess in advance how a patient is going to react to a meal or to a bolus of insulin. So we do not have any foreknowledge when we simulate of what's going to happen. We set the scenario. The patient is going to eat 60 grams of carbohydrates at one hour and take a bolus of three units at the same time. But we really don't know what the -- how glucose is going to evolve in the next six hours. We have to actually run the software to figure it out.

Once you've created an insilico subject, you get to create an in silico population, and that's basically to have as -- about 100 subjects, in our case, that span all the variability that is observed

in vivo. And that basically allows us to have very different subjects, the same way you observe them in your clinics.

We finally validated with external datasets both from hyperinsulemic clamp and from the direct net study. We also did some validation using basal rates and carbohydrate ratios and different measures of insulin sensitivity. And I have to thank Dr. Buckingham and Dr. Clarke for sharing their -- the data of their patient in these particular case that allow us to verify that the patients that were present in that simulator had characteristics that corresponded to what's observed in vivo.

Finally, that simulator was accepted in January 2008 by the FDA to replace pre-clinical studies in closed loop trials. And so it means that it was at least good enough to replace animal trials, and it has been at the foundation of all of our investigational device exemption for such studies since.

The current in silico population is made of 300 patients -- 100 children, 100 teenagers, and 100

adults, all Type I. They can be admitted to the CRC. They can be tested in different ways to extract information about them, and of course you can submit any scenarios composed of any combination of meals, carbohydrate intake of any form, and insulin injection of any form. And that's the demographic characteristic of the population. All right.

Now, enough about the tool. Let's talk a little bit about the study that we did. So we used this simulator to study the effect of how tight that ISO standard that we've been hearing about, what's the effect of its amplitude on glucose control. And so the only thing that I've done here is reproduced that ISO standard, you can see that, of 275 milligram per deciliter. You have a fixed error, and after 75 milligram per deciliter, you have a relative error. In this case, it was 20 percent plus/minus 15. I wanted to emphasize with another figure that what we're talking about is the famous 95 percent, and not all the points have to be in there. And so the color-coded graph that is on the right here, the more red you are, the more data points you're going to see. The

more blue you are, the less data points you're going to see, but you're still going to see data points in the blue area.

The first study we did was to look at the detection of hypoglycemia based on different levels of sensor accuracy. And so that's been shown to be a very important feature to Type I diabetic patients, is that fear of hypoglycemia is what generally limits their capability of controlling their (inaudible) their glucose.

Now, if you take the actual sensor -- well, the actual ISO standards of 20 percent, you can see that first, if your true sugar is 70, there is a 50 percent chance of detecting or not. So that basically means that what we're studying are unbiased sensors. We're only looking at the spread as defined by the ISO standard.

Now, when you go down in true plasma glucose, of course, your probability of missing the hypoglycemic event diminishes, and you can see that about -- if the true glucose is about 60, you have about a 10 percent chance of missing it. When it's

about 50, you barely have a one percent chance of missing the event, and that was demonstrated in a talk earlier.

Now, if your accuracy becomes better, so if you move from 20 to 15 to 10 and to 5, you can see that your probability of missing a hypoglycemic event is reduced dramatically. And I'm going to be particularly interested in the .60, and you can see that if your true plasma glucose is 60 milligrams per deciliter, with a five percent errors -- of course, if you have a perfect sensor, you're not going to miss the hypoglycemic event. At a five percent accuracy level, there is almost no chance of missing that event, either. It's .001 or something like that.

At ten percent, you have about one percent chance of missing a hypoglycemic event, which is reasonable. And you can see that between ten and 15, something happened that actually makes that risk of missing an event increase quite dramatically. And the difference between 10 percent accuracy and 20 percent accuracy is actually a factor 10.

Now, we also looked, in a second study, at

what the effect of accuracy is on one very simple treatment, treatment of hyperglycemia. So we used the previously described simulator and the previously described model of SMBG noise, or SMBG error. Each in silico patient starts the experiments stable at 200 milligrams per deciliter, and for each of these patients we designed a perfect bolus that would bring the patient down to 100 within four hours. And so basically you measure glucose at times zero, that measure is perfect, and you have the perfect bolus that will bring the patient down to 100.

And then we repeated that experiment with actual measure designed with the SMBG error model and looked at what was happening. Now, of course if you measure lower, you're going to treat a little lower and you're not going to achieve your target. If you measure higher, you're going to actually see that you go lower than your target, and you might actually get into hypoglycemia.

Now, this presented the result that we have from this experiment, and you can see that at five percent error, I mean the minimum that we saw was

about 90. The maximum that we saw at about 110. So by all practical measures, they achieved perfect treatment of hyperglycemia. And you can see how the capacity of treating hyperglycemia deteriorates with increased errors, SMBG errors. At 15 percent, you actually start seeing occurrence of hypoglycemia. At 20 percent, you see both hypoglycemia and for some subjects, they did not even reach the 140 milligram per deciliter limit. So they probably have to treat further down the line.

And these results are presented in bar graph. You can see that the risk of hypoglycemia is zero percent up to 10 percent of error, and then rises to 3.5, and then 5.5 percent.

We also looked at glucose variability, and this time we actually simulated a meal. So we measure at the beginning of a meal for the meal bolus, and we also have a measure two hours later. The meal bolus was intentionally underestimated so that we would be high two hours later. And then the patient treats from the point there are two hours later.

If you under-measure at times zero right

before the meal, you're probably going to under-bolus. You're going to get a little higher. Let's say you over-measure at two hours later, you're going to over-treat, and then this time probably end up lower than you initially intended. The results that are presented here are -- it's called the control variability grid. On the X axis is the lower points that you attained. On the Y axis is the highest points that you attained. So basically, if I was an absolutely perfect control, I would be somewhere around this in the A zone. You will note the similarity with the Clare Error Grid. Anywhere around that arc is proper treatment in the D zone, and then these are the error zones of danger. For example, if you're in the E zone, not only do you go over 300 milligram per deciliter after your meal, but you went hypoglycemic afterwards.

So you take, basically your glucose curves. You obtain the maximum to minimum. Take these two numbers, and that gives you a point on the grid. And you can see the difference between white and black. White is almost no noise; black is a lot of noise. You

can see that when you had an inaccurate sensor, you start dealing with dangerous situations.

Finally, we studied the long-term effect of SMBG accuracy, and so it's -- it was a similar simulation, but with three meals a day for ten days for each patient. And what we observed first in a nominal -- we fixed the risk of hyperglycemia for each patient at 15 percent, which is basically one day out of -- one day every week where they would experience a hypoglycemic event.

And when we augmented the sensor errors, we actually observed that these hypoglycemic events became more frequent. And so we estimated that each patient had their own -- their own aversion for hypoglycemia, and so we dialed back in all the traces for each patient to their original nominal aversion of hypoglycemia, which is the 50 percent I just talked about, and that dial back in, we transform into HbA1c using the ADA formula. And that basically shows that at nominal level, it's exactly at 15 percent. Five percent error, you had no more hypoglycemia. And then 10 percent, 15 percent, 20 percent, you can see a

linear increase. In transforming into HbA1c, you can see a moderate increase from nominal to 20 percent, of about .4 percent of HbA1c.

So in conclusion, in silico experiments allow for fast and inexpensive study of clinical consequences. We've shown that SMBG accuracy has dramatic consequences in terms of hypoglycemia detection. We've shown that it has also very important consequences in terms of treatment of hypoglycemia. It also augments glucose variability, due to the treatment of the meal. And it had a moderate effect in HbA1c over a long term period.

I would like to end up with very wise words, which is that all models are wrong, but some of them are useful.

(Applause) I would like ....

DR. GINSBERG: Ginsberg, New Jersey. Marc, very nice study.

When you do a model of merits, it's important that you consider all the system merits. When you modeled it, what did you use as the error for carb counting, for error in the constants, and error

for the absorption of insulin?

DR. BRETON: So what I really -- so there's different answers for the different parts of your question.

So first, the error in carb counting, I was not necessarily interested in that. What I wanted to see was the sole effect of an error of glucose measurement. So of course I can induce some error of carb counting, but that would have induced errors in my -- in my results that would not have been caused by the SMBG. So there was no errors in carb counting.

Now, for the difference in insulin absorption, as you very well know, each of our subjects has a specific insulin absorption, and it's not going to change during the course of the day. What we are simulating, the observed variability, is not within one subject, but by having different subjects with very different characteristics. And so even though our subjects keep the same insulin absorption throughout the day, Subject Number 2 is going to have a very different insulin absorption than Subject Number 1, and that's how we account for that

variability.

DR. GINSBERG: Okay. I'll present some of that in my talk. But when you put in all the errors, would you find, for a patient with Type I, based upon a meal dosing, that the error in glucose is about four to six percent of the total error in the system? Doubling the error in glucose goes to almost nothing.

DR. BRETON: It's not surprising. As Dr. Clarke was telling me earlier before this talk, if a patient forgets -- correct me if I'm wrong -- if the patient forgets one insulin shot in one week, their HbA1c is probably going to rise by the equivalent amount of what I showed today.

DR. CLARKE: One final quick question.

DR. BRETON: Okay.

DR. HARPER: I'd like to remind people with questions to please state your affiliation, as well.

MR. HUANG: Okay.

DR. CLARKE: Even if you've asked before.

MR. HUANG: Dijia Huang with Bayer Diabetes Care from Indiana. It's just a quick comment.

And I feel generally people use simulations

for two purposes. One is to identify the trend, such as the percentage that we miss the hypo event versus the accuracy of the meter. The second application is to read exact number, such as, under what accuracy of the meter - - meter accuracy and what's the exact percentage of a missed hypoglycemic event?

And personally, I feel for the first application, for the trend, simulation is an excellent tool. But then when we go to -- when we go beyond this trend, go to the -- to read an exact number, to reach the conclusion, use the exact number, then I think it becomes weak. My concern is, such as the differential equation you showed, and in it, minor adjustment on those constants could shift the curve left and right and make the -- make us get different readings.

Just my short comments.

MARC BRETON: I'm going to have to both -- which is interesting -- both agree absolutely with you and disagree. First, let's go on with the perfectly agree. Simulations is a limited tool, and it had -- it is designed for specific goals, and generally it

shouldn't be used outside of these goals. You will note also that I only claim to replace pre-clinical data. At no point in this talk will I ever claim to replace clinical studies.

The place where I disagree with you is you're saying any shift of these constants in the equations would shift the curve left or right. Well, you're right if I was using only one set of such parameters. Now, this particular simulator has the advantage of having a 300 subject population, in which we claim we actually represent the variability. Actually, we represent more than the variability that is observed in vivo. So that particular shift that you're talking about, we actually see it here, and we account for it. And the result that I present to you, of course, averages over all these subjects.

MR. HUANG: Thank you.

DR. CLARKE: Thank you, Marc.

(Applause) Our next speaker this morning is Dr. Stephen Brotman, who is vice-president of payment and Healthcare Delivery Policy at AdvaMed. Dr. Brotman, like Dr. Shuren, has both an M.D. and a J.D.,

and he's Board- certified in pathology and has worked as a senior regulatory and research attorney at Wyeth, which is now Pfizer this week, and who knows what next week.

And Dr. Brotman's going to talk about tighter performance criteria for blood glucose meters and whether or not they're needed.

DR. BROTMAN: Thank you. Good morning, good afternoon maybe. My name is Steve Brotman, and I'm currently a senior vice president at AdvaMed, which is the Advance Medical Technology Association. I'll be discussing this morning the industry perspective on tighter accuracy requirements for blood glucose meters.

This is an incredibly important issue for which industry has been engaged for years in development of the standard and ongoing advances in technology to bring the latest innovations in blood glucose meters to patients. We thank FDA for the opportunity to speak today at this meeting.

Just a little bit of background about our organization. AdvaMed is the world's largest

association representing manufacturers of medical devices, diagnostic products, and medical information systems. Our member companies produce the medical devices, diagnostic products, and health information systems that transform - - they're transforming healthcare through early disease detection, less invasive procedures, and more effective treatments.

The medical technology industry is strongly committed to designing and manufacturing blood glucose meters that meet the needs of individuals with diabetes. We share the goal of improving meter accuracy.

Meter accuracy includes not only analytical performance, but the key areas impacting blood glucose meter accuracy include use error and interferences. The industry has made tremendous strides in improvements in both reducing use error and reducing the impact of interferents. Both are incredibly important aspects of use of blood glucose meters by self-testers.

The standard currently governing blood glucose meters for self-testing, ISO 15197, itself

recognizes the importance of usability improvement. Specifically, it notes that the goals for performance criteria should be weighed against the capabilities of current self- monitoring devices. Furthermore, the standard notes that care should be taken implementing performance requirements that cause manufacturers to focus design improvements on analytical performance at the expense of other important attributes, such as greater convenience and greater compliance. Thus, the standard acknowledges the careful balance of these factors and the minimum acceptable device performance for glucose meters for self-testing. The standard supports performance improvements beyond analytical performance, such as advances that reduce dependence on user technology, otherwise referred to as patient usability.

Industry has made major advances through usability improvements since the approval of ISO 15197. These address ergonomic and human factors that are incredibly important to patients who are self-testing. Manufacturers also integrate overall usability engineering in device design. There has

never been an array of usability -- there have been an array of usability improvements. Such features will, of course, vary per meter. No single blood glucose meter will meet the needs of all patients.

While I'm not going to go into detail on each and all the various usability improvements, it should be noted that there are numerous improvements for which we have outlined. This gives a sense of the types of advances made by industry which contribute to the overall improved device performance. Among others, they include faster test times; smaller blood samples for decreased comfort (sic), and I just want to mention under that, patient discomfort was, in fact, a focus of the discussion in October 2001 FDA panel, where the panel encouraged FDA to approve blood glucose monitors with alternate site testing labeling instructions. Although alternative site testing may introduce more inaccuracy in precision, the panel cited that reducing pain would improve overall patient compliance with their testing program.

If we look at additional usability improvements, they also include overall increased

robustness, such as cleaning solutions and hardware testing, enhanced meter displays, such as larger readouts, back lighting, and universal symbols. They include ergonomic design, such as buttons and meter size, including smaller meters for more discreet testing.

Another usability improvement would be non-changeable unit of measure by the user. And in some cases, no coding or calibration or timing is needed.

Other usability improvements include wider temperature range; improved range and stability for longer use life and decrease susceptibility to exposure; biosensor in addition to photometric technology; plasma- referenced results; integrated meter and lancing devices; improved voice simulators for the visually impaired; flagging test results, for example, meal markers; innovative software to organize meter data; and in some cases, no individual test strip for lancet handling, to reduce use error and to increase the ease of use.

And the list goes on. I will not go through all the improvements, but they are quite significant

and certainly have impacted increased compliance and frequency of patient testing.

Just a word about additional usability improvements which are listed here. I'm not going to go into these, but it's evident there are clearly many options available, reflecting significant technology advances over the years.

Most of these usability improvements are actually beyond the current standard. They are not required under ISO 15197. They are the result of manufacturers' commitment to constant innovation to meet the needs of the self-tester. It should be noted that beyond the usability improvements, there are other significant improvements, including advances that have reduced susceptibility to interference, such as hematocrit. All of the advances are part of industry's efforts to support technology innovations that improve health care and lead to better outcomes.

So the question comes, where are we today? Well, multiple clinical studies have shown that home blood glucose meters meeting the current standard for accuracy, which is ISO 15197, are associated with

improved glycemic control and produce positive clinical outcomes in such large randomized clinical trials, such as DCCT, the Diabetes Control and Complications Trial, which looked at the effect on insulin-dependent diabetes mellitus.

Furthermore, the current standard has been shown to produce clinically acceptable results. This was illustrated in consensus error grid analysis which has been constructed as an unbiased tool to analyze the clinical significance of blood glucose self-test measurement errors.

According to the consensus error grid analysis, 96 percent of the results fall within the range indicating no effect on clinical action, and the remainder fall in a range indicating altered clinical action with little or no effect on clinical outcome. Error consensus grid is a tool, we think, that is strongly considered an important tool for objective outcome assessment for blood glucose devices for self-testing.

We also believe that the impact of criteria changes on current meters should be considered in any

change to accuracy. This is important, because some improvements will be difficult to sustain. Potential negative impact may be increased test time, increased sample size; narrow operational conditions, such as temperature or relative humidity; potential bulkier meter; and impact on usability and how overall user-friendly a meter may be. Cost is also a factor for patients in terms of affordable meters that meet the needs for self-testing.

It is important to recognize that the current accuracy criteria in the ISO standard is specific to self-monitoring use, not professional use. Industry fully supports efforts to develop increased accuracy requirements for blood glucose meters used in hospitals and long-term facilities. A standard is being completed to address this issue as we speak through the Clinical Laboratories Standards Institutes, CLSI, through updating of POCT(12), otherwise referred to as point of care blood glucose testing in acute and chronic care facilities. FDA, industry, and key representatives of the clinician community are all engaged in this effort. The

political consequence of an inaccurate glucose result in these particular settings, which include immediate treatment decisions and generally increase susceptibility of the hospitalized patient merit specific guideline development for meters used in those settings.

We fully support those efforts. We also acknowledge that tight glycemic control is an important issue for hospitalized patients and the clinicians who are treating them. Importantly, we note that the current standard, which is for home use, was never intended for patients on tight glycemic control in the hospital setting. It will be addressed in POCT(12) and we fully support these efforts to address this in POCT(12).

As previously discussed, there has been a number of recent advances and evolution in blood glucose meter innovation by industry, contributing to an enhanced overall device performance. As part of that commitment to blood glucose meter innovation, industry is actively participating to update ISO 15197. In addition to other revisions to further

improve the standard, industry strong supports revising the accuracy standard in section 7 of ISO 15197 for system accuracy of more than or equal to 95 percent of the individual glucose results to be within plus or minus 15 milligrams per deciliter from the reference glucose values, and within plus or minus 15 percent. At present, the current standard is plus or minus 20 percent.

There is also recognition by the standard review working group that the standards for hospital is CLSI POCT(12), and not ISO 15197. In addition, industry is actively engaged in discussing how to best deal with interfering factors in the standard. Industry is strong supportive of all these efforts, which we expect to culminate in a newly updated standard of great importance to the blood glucose meter industry and the patients using these innovative technologies.

Industry has also been proactively engaged, as mentioned, in an update of POCT(12). It should be noted that the document that the document is near completion and will set out the latest guidelines for

blood glucose monitors in the hospital setting, including considerations of overall accuracy and tight glycemic control in that setting.

We also believe there are other mechanisms to support appropriate use of standards. FDA could formally adopt the ISO standard into FDA guidance to support enforcements for use of standards for self-testing only. But we continue to encourage that usability and analytical performance should be carefully weighed to assure meters that meet patients' needs.

Thank you for the opportunity to present today. Industry looks forward to ongoing work with FDA and the global standards community in systemic review of ISO 15197, as well as the completion of POCT(12) to support blood glucose meter innovation and the needs of users. I'll take any questions.

(Applause)

DR. CLARKE: I have one question. I guess maybe I'm -- I'm inferring, from what you've said, that industry is trying to separate or would like to perhaps separate out home use by patients from

intensive care use. And if that is so and they would like to have an ISO standard to improve the home use, but not necessarily the ICU use, I'm sure that industry wouldn't want to forego those 250,000 blood tests that are done in the ICU at Barnes Hospital every year, or -- I mean, I'm -- I don't get that -- seems like that's a disconnect.

DR. BROTMAN: Well, I -- I see your concern, and I think the best way to address it is that there are members in Panel 2 from the blood glucose working group that have been talking about this for quite some time in quite some detail. And I think it's probably better addressed during that panel, if you wouldn't mind.

Yes, ma'am?

MS. RUTHERFORD: My name is Diane Rutherford, with Kenbuck(ph) Consulting in Dallas, Texas. And as a representative of industry, I thought this might be better suited for you.

We're talking about increasing the capabilities, the accuracy of the monitoring devices. How does that affect the accuracy of the dosing

instruments? Your syringe dose tolerances are pretty wide, from my experience in industry. And is it fair to make the monitors more accurate and not make the dosing requirements more accurate, as well?

DR. BROTMAN: Well, I -- I think that that's a valid point for discussion. I think that's a discussion that's probably an ongoing discussion, and it's probably something that is continuing at this time. Again, I think that's something that could be addressed, and probably has been addressed, by the blood glucose working group, and you may want to pose that question.

DR. CLARKE: One further question? It's a long way down front.

MS. SOLDO: Sorry.

DR. CLARKE: It's okay.

MS. SOLDO: Monnett Soldo from OptiScan in California. We are not participants in the industry group that you mentioned, as a small startup. I was wondering if you could share with everybody, since you said the new proposed CLSI standard is near completion, what exactly that standard would be?

DR. BROTMAN: Again, I think in terms of all the details that have been worked through and considered, the members of the working group would be the best members to be able to address something like that.

MS. SOLDO: Okay. I was thinking at the level of, you know, we're talking plus or minus 15 -- plus or minus 15 percent. Can you summarize at that level?

DR. BROTMAN: Again, ask me specifically your question? The POCT(12)?

MS. SOLDO: Yeah.

DR. BROTMAN: I think that's a question that's probably best aimed at the working group, if you wouldn't mind.

DR. CLARKE: Other questions? Thank you very much.

DR. BROTMAN: Thank you.

DR. CLARKE: I'm going to just wing it, Barry.

Our next speaker is Dr. Barry Ginsberg, who is the CEO of Diabetes Technology Consultants. Dr.

Ginsberg is former vice president of WorldWide Medical Affairs for the Diabetes Division of Becht & Dickinson, where he led the diabetes program there for about 17 years. He is a former professor of internal medicine at the University of Iowa. He has done his training at Beth Israel Hospital, as well as at the diabetes branch at the NIH.

He has been a consultant to the technological aspects of diabetes for a number of years, and he's going to talk to us today on Industry Perspective: Tighter Performance Criteria are Achievable and Appropriate.

DR. GINSBERG: Thank you, Bill.

I actually had tried to make it easier for Bill by actually putting a small CV up in here, but ....

Let me talk a little about my conflict, first. I'm a consulting medical director for Agamatrix, which is a blood glucose monitoring company, as well as a speaker for Bayer and a consultant to the Juvenile Diabetes Research Foundation on the Artificial Pancreas Project.

Indirectly, I work for companies who've worked for Roche and for LifeScan.

Let me talk a little about the industry perspective, tighter performance control -- criteria are achievable and appropriate. I think I'm the first speaker up here today who has an academic, a clinical, and an industry background. And all three of those go into my opinions today, which are solely my own.

As an overview, we'll produce some background. We'll talk about measurements of inaccuracy, necessary accuracy, and a little bit of outliers, and something about the sources of error. I'll talk a little about the current technology, where the accuracy is, and methods to improve it. I'll talk about testing and reporting, whether the testing ought to be done by the companies themselves internally or externally by a notified body like company, and whether testing should be done initially or periodically. And then finally, since this is a consumer product, what we ought to do about consumer labeling. And I think I can do that all in 20 minutes.

All right. When we talk about accuracy, we're really talking about accuracy and precision, and you really need both. A device which looks like this in terms of its accuracy, when you average out that accuracy, it's actually perfectly accurate, even though none of the individual values are even close to being accurate. So average accuracy by itself doesn't necessarily help us.

Similarly, average precision, by itself, doesn't help, because that's pretty tight and precise but nowhere near the value that we're interested in. And it's only when you consider both accuracy and precision that you get values which are right at your target. We'd like that with one number. And the number that comes closest to that is the number that's been given a couple of different names today, or mean, absolute, relative error, mean average deviation, relative average deviation, and so on.

And what it really does is, you just look at accuracy. What you're looking at here, if this is plus ten, this is minus ten, that's plus ten, that's minus ten. When you average those all out, that comes

out to zero error. When you do this, you take the absolute value. So a plus ten becomes a plus ten, a minus ten becomes a plus ten, plus ten and minus ten again -- the average error here is ten percent. And so that single value, you're taking the absolute value, gives you an error which is useful in this. We also need to consider bias.

Now, we've talked a lot about ISO 15197. It's an international standard. It contains both clinical and laboratory standards, as well as other parts. Most importantly, it contains the study design of how this study should be done. We've talked about values, that 95 percent of the values should be within 20 percent of the reference value for values greater than 75, and 15 milligrams per deciliter for values less than 75. And shown on a graph, that's sort of what this looks like.

Within the ISO, there's also an extended ISO standard, which is not required for approval, but is suggested it be done as well. And the extent that ISO looks at accuracy of not only 95 percent -- 20 percent, but at 15 percent, 10 percent, and 5 percent,

and that's shown in the graph here in the orange, yellow, light green, and dark green.

More importantly, when you ask the question of, what do these do in terms of hypoglycemia? Because I believe that glucose monitoring is most useful at a patient level. Let me point out here, I am only talking about self-glucose monitoring here. I am not talking about hospital use of meters in any part of this talk. It's really useful in two respects for the patient. One is it's useful in helping them select the dose of insulin appropriate at a meal, and secondly, it's helpful in letting them know when they're becoming hyperglycemic. And we'll talk about those two separately.

But when you look at the ISO standard and ask, So if the actual blood glucose is 70, at the various ISO standards, what is my 95 percent confidence limits on what I'm going to see? And so at the ISO standard of 20 percent, that 70 is somewhere between 55 and 85. Well, if I'm 70 and the meter's telling me 85, I'm not sure I'm getting the appropriate information.

At 15 percent, it goes up to 81; at 10 percent, 78; and at 5 percent, it's down to 74. So the more accuracy here is actually probably pretty important.

ON the other hand, when you consider the effect of blood glucose in terms of selecting an insulin dose, I think it becomes pretty unimportant. Let's consider all the errors that go into selecting an insulin dose. So I'm sitting down to a meal, and the first thing I'm going to do is count my carbohydrates. Well, how much of an error is there when I figure out how much food is in that meal? Well, it turns out the likely number is 15 to 20 percent. Some people are significantly less accurate than that and may have numbers up to 25, even 35 percent.

So I now have a number, I'm going to eat 60 grams of carbohydrate. Well, I'm going to take that and divide that by my carbs to insulin ratio, which is ten, saying I need six units of insulin. How much error is that constant? Because it turns out that most patients are given a number when they develop diabetes, and no one ever changes that number again.

They have that number for the rest of their life. The likely error there is at least ten percent, and maybe as high as 25 percent.

And then finally, I actually get my calculation, I need six units of insulin, and I inject it. How much error is there in that six units of insulin being absorbed? And again, it depends on the kind of insulin. For a rapid-acting insulin like lispro, it turns out that's about 20 percent. For a less-rapid acting insulin, like regular insulin, it turns out that's about 35 percent.

Now, when you have a series of errors -- engineers like to stack them -- but it turns out the correct thing to do is to take the square root of the sum of the squares. And when you do that, you find out that the total error here is about 27 percent, which gives you a 95 percent confidence in the error of plus or minus almost 70 percent. And likely error here is 46 percent, with a 95 percent confidence limit on that of plus or minus 115 percent.

Now, it turns out -- I didn't put blood glucose monitoring in here, because it turns out that

if you take blood glucose monitoring and change the error, the average error, from five percent to ten percent, you barely change that number. You increase it by one to three percent. Well, at three percent at 27 percent, I increase my error by only ten percent. And here, even less.

So it turns out that blood glucose monitoring error, in terms of selecting a meal, is actually probably relatively unimportant. ON the other hand, as you look at Number 4, blood glucose monitoring error when I select -- when I want to know if I'm hypoglycemic, is actually pretty important.

Now, based upon that, and this is only personal opinion, what I would suggest is how much accuracy you need depends upon who you are. If I'm a patient with Type II Diabetes on diet or non-hypoglycemic oral agents, I don't need to know about hypoglycemia. I'm never going to become hypoglycemic. There's nothing there to make me hypoglycemic. And so the current standard of 20 percent is actually fine.

If I'm a patient with Type II Diabetes on insulin or an oral agent that does cause hypoglycemia,

which is becoming rarer and rarer, then I probably need a little more accuracy, and 15 percent is probably appropriate for me.

If I am a patient with Type I Diabetes who is now running my blood glucose an average of 80 or 90 or maybe 100, now accuracy in terms of hypoglycemia becomes very important for me, and an average accuracy of ten percent would actually be very helpful to me. I might comment here that laboratory accuracy at two percent would put me up here.

Just an interesting point as an aside here. When you start to consider the accuracy of these meters, you start to consider how you're going to test them. If I have an average meter with ten percent -- 95 percent confidence limits or a four percent average in accuracy, I can't test that with a laboratory instrument. Because laboratory instruments have two and a half to three percent inaccuracy. That means the error in my testing instrument is going to be 30 to 40 percent of my total error. Then when I get down to this, I'm going to have to start going to GCMS spec or another really tight reference value, which is

going to be very hard to do.

Now, we talked a little about error grids. And what I've shown here is something I consider very important in terms of outliers. This is not real data, because the real data it's based upon is owned by a company that I couldn't -- I actually didn't have a chance to ask them to use the data. So this is data that I've made up, but it does reflect real data. So it's for illustrative purpose only.

And here you have 5,000 samples on a consensus error grid. It looks pretty good, except there are five values there, or about .1 percent, which are not so good. It's a value of 290, which is showing up as 525; a value of 230, showing as 475; a value of 205, showing up as 350. At .1 percent, the average patient with Type I Diabetes will get at least one of these a year, and probably more. And if they don't -- if they believe this, that's going to lead to serious hypoglycemia when they take that insulin for that.

So I think outliers are important. I don't know what to do about them, but right now we're doing

nothing about them. We're not even talking about them. I think it's important that we start talking about them.

Now, where are the sources of error that's we're going to deal with? Well, the first is intrinsic sources. Manufacturing variation -- there are small variations in the size of the well or the size of the silkscreen-printed electrodes. There are differences in how well the enzyme is laid down in that well or print screen. There is the age of the strip, and mediator oxidation, all of which add to intrinsic error of the strip.

This physical location -- temperature can affect this, and altitude can affect this. There are interfering substances, and I won't go into these in detail, because they have been before -- intrinsic urate triglycerides for some oxygen, for glucose oxidase, ascorbate, acetaminophen, L-Dopa, Enflazamide, and for glucose, the hydrogenous PQQ maltose, icodextrin, and xylose.

And then finally, there are patient factors. From a medical point of view, there's hematocrit. From

an educational point of view, there's technique, coding, and hand-washing, all of which add to the errors that we're dealing with.

Now, this is a paper by Freckmann from Diabetes Therapeutics and Technology in December. And God bless him, he actually took a whole bunch of meters and tested them all in the lab, and actually found that for the most part, they were pretty accurate. That you can see all of them met the 20 percent standard, many of them met the 15 percent standard. As you go down to the ten and five, fewer and fewer of them did.

If you look at data from FDA submissions or in package inserts, you find data here, and this is from nine companies which we could get this from, and you can see that for the most part, they all met the 20 percent. A number of them met the 15 percent; some ten, and some five. Some of these are clinical and some of these are laboratory, but Company E, actually, I know is clinical. And they did pretty well. They can meet a 15 percent standard. They can almost meet a ten percent standard. The question you could ask is,

do you need that? We'll get back to that in a second.

Now, there are methods to improve this. Hematocrit can be measured by AC impedance or by dynamic electrochemistry. Temperature and altitude can be met by thermistors, and again, dynamic electrochemistry. Manufacturing variation, QC, interference by multiple electrodes or dynamic electrochemistry. Coding, by no coding meters, either by code selection or coding on the strip. And finally, aging -- some of the meters actually, they use cartridges, print the age on the cartridge, and the meter won't work if the strip is too old. So there are approaches which will get us more accurate as we go further along.

Now, a number of years ago, I made a statement you can trust device manufacturers in clinical trials. If it's EU certified, you can ask the notified body testing, and you can pass it. I've seen some data recently which makes me a little less confident about manufacturing testing, particularly from some of the smaller manufacturers, and I sort of wonder about that. So I'm going to ask the question,

Is it appropriate now to go beyond internal testing by the manufacturers and go to a system closer to Europe, where you have notified bodies or other groups that actually do the testing for you, so you can be sure that when you publish the results of your accuracy, it's the same for everybody?

Blood glucose meters have become consumer products. When companies advertise on blood glucose monitors, they talk about size. They talk about shape, they talk about color. They talk about convenience. Those are consumer features. Some of them are talking about accuracy, but for the most part, we're talking about consumer products, in which consumers are making decisions about the device they choose. And for a consumer product, information is king. You need to create informed consumers, and the consumer needs to have standardized, accurate information.

Because of that, I think we ought to go to external testing. It's a standard protocol with clinical testing of random lots, not specific lots. A standardized analysis of the data in order to produce

standardized labeling. And put the label right on the strip container, saying exactly how accurate how you are in a method that anyone can understand. As a matter of fact, at a fifth grade reading level -- as a matter of fact, if we come up with a better number, something akin to a batting average, I think that will be even better.

I put this in gold because I don't know whether we ought to have -- we ought to develop a standard protocol for outlier analysis, but I don't know what to do with it after you get that.

Clinical testing ought to be periodic. It ought to be every six to 12 months so we know that the accuracy of the initial test is continuing. It ought to be done on random lots. I think reasonable failures don't require a corrected -- they require correction, they don't require recall. And outlier testing ought to be continuous, but outlier testing requires so many strips. You need 10, 20, 30, 50,000 strips to test for outliers. I think that has to be done by the manufacturer.

So my recommendation is, then, our accuracy,

the substantial majority of users, which is those patients with Type II Diabetes on diet or non-hypoglycemic oral agents, don't require more accuracy than we have today. And if we get -- if we demand more accurate devices, we're going to lose other consumer features and/or make them more expensive. So I think we ought to keep the current ISO standard as the minimal acceptable clinical accuracy, although it ought to be done with random lots.

I think we ought to be labeling devices with what the true accuracy is. The strip container should be labeled with the mean average rule of error, and the ISO errors, and maybe, although I'm not sure, labeled as group-appropriate. This is appropriate for Type II's on oral agents; this is appropriate for Type I's on insulin, et cetera.

Testing ought to be done initially with random lots, external control at a notified body-like device, and periodic testing of random lots and better testing of outliers.

Labeling -- I just put a hypothetical label here -- should show, develop a standard label, sort of

like nutrition labeling, or Energy Star labeling. I mean, all other consumer products have this. Why shouldn't this consumer product have it? So develop a standard label so you can look at it and know exactly where the average error is going to be, where the extended ISO error is going to be. Put in MARE, or even better, a batting average. Put in extended ISO data, and question about intended consumers.

So in summary, then, I think blood glucose monitoring is currently accurate enough for a substantial majority of patients. Now let me say, this is not strips. Those patients do not use the majority of the strips, but it is the majority of patients.

Better accuracy is achievable, but not necessary for everybody. Since it's a consumer product, better labeling is essential. We ought to do that with external testing with a standard protocol and periodic testing. And a standard label, like a nutrition label.

With that, I'll stop, and ask if there are any questions.

(Applause)

DR. CLARKE: Are there questions for Dr. Ginsberg? I have a question.

DR. GINSBERG: Okay.

DR. CLARKE: You talked about labeling, and you talked about pitching it at a fifth grade level, which I think is really a good idea. But it seemed to me like the labeling that you were talking putting on the strip bottle could not be interpreted by somebody with a fifth grade or an eighth grade education.

DR. GINSBERG: I'm not sure that -- I'm not sure that's true. If you present a booklet to go along with this -- when you go to Consumer Reports and look at their red circles, black circles, and half circles, they don't explain to you what it means. And it's not that hard to say a MARE, the lower the number it is, the more accurate it is; that for the ISO standards, the higher the number it is the better it is. And that for different groups, you may require different amounts. And we have to agree on what that is, but I think that could be done.

DR. CLARKE: Go ahead, please.

MR. BASHAN: Eran Bashan from Michigan. I actually think that your data suggests that we're much better positioned than we think, because you know, even if you look among insulin testers, 60 percent of the folks on insulin use either long-acting insulin or premixed biphasic insulin. And their dose doesn't actually depend on the glucose level. And if you look at, you know, your major concern, the outliers, there are roughly ten billion tests done in the U.S. every year. If you think that one percent of them are extremely inaccurate, that's 100 million tests per year.

DR. GINSBERG: Point one percent.

MR. BASHAN: If you look -- .1 percent, ten million. If you look at, you know, mortality from diabetes, that's two orders of magnitude less than that. So you can actually claim that 99.95 percent of the tests done today are extremely safe.

DR. GINSBERG: But there are 50,000 serious hypoglycemic events per year.

MR. BASHAN: I know. But that suggests, given the amount of measurements, that the real

accuracy that you have today is much better than the ones actually reported.

DR. GINSBERG: Thank you.

DR. WHITE: I'm Neil White. I'm a pediatric endocrinologist from Washington University in St. Louis. I think that education of the consumer is really important, and you've made -- you've made that point. And we talk about error in here, I think some of us in the audience are learning a lot about error in the measurements. But I don't think our patients, especially the mothers of our children with diabetes, understand error at all. They expect the number to be a number; okay? They don't see -- they call us when they do two measurements in rapid succession and one is 120 and one is 135, okay, which we all know is within the error of a test. But they don't understand that. There's a huge opportunity for education for the consumers understanding what these strips are capable of doing and what the results mean.

DR. GINSBERG: I agree. As long as Neil is up and reminding me, actually, let me actually point out in terms of accuracy that a number of people have

talked about the DCCT. The average error of the meters in DCCT was between 13 and 15 percent in the hands of patients, significantly higher than we're talking about today, by a factor of three.

DR. CLARKE: Other questions? Thank you.

(Applause) We are going to break now for lunch until 1:30. At that time, we will reassemble here, and a larger, expanded panel will be present for you to question. So while you're having lunch, ask your friends, or you're not-friends, what they thought of this morning's presentations and what questions they had, and let's make sure our questions are answered at 1:30. Thank you.

SESSION 1 PANEL DISCUSSION

DR. CLARKE: Please take your seats. We are ready to begin this afternoon's session. We are about to have a panel discussion, and the panel includes all of the speakers from this morning, plus some additional distinguished individuals, including Ellen Ullman, who is a diabetes advocate and Vice President of Children with Diabetes, and works closely with Kelly Close and Closer Concerns and is a parent. And

so she is really the Type I Diabetes advocate here, and we're going to all want to hear from her.

Dr. David Klonoff, who is sitting to her left, but to my right, who is Clinical Professor of Medicine at UC San Francisco, the founder of the Diabetes Technology Center, the Editor-in-Chief of the Journal of Diabetes Science & Technology, and the Medical Director of the Mills-Peninsula Diabetes Research Institute, in his spare time.

And next to him is Dr. Alberto Gutierrez, who is the Office director or Director of the Office of In Vitro Diagnostics at the FDA.

And this is going to be an exciting time. I would stress to the people here, when a question is asked to you, make sure you push your microphone until the red light comes on, and when you're finished talking, make sure you turn it off, because otherwise people will hear what you say.

(Laughter) And it could be embarrassing.

And at -- at this time, I'd like to invite people to come to the microphone, and you may ask any question that you wish. Don't be shy. There was

someone who stood in line a long time and didn't get to ask a question.

Okay. I'm going to start out with a question for Ms. Ullman. Because you've been through this business of testing blood glucose a lot in a child, and because I'm a pediatrician I can empathize with that. And I guess I'm wondering, do you feel that you needed a blood glucose monitor for your son which was significantly more accurate than you had? And -- here's the second part of the question -- would you be willing to give up something in order to have more accuracy? And that giving up might mean taking longer to get a test result, giving a -- maybe having a larger drop of blood, or something else?

MS. ULLMAN: Oh, absolutely. We needed -- my son was one when he was diagnosed. He's now 22, so we've done 21 years of blood glucose testing. And when you have a little child, you know, one year old, and you see a 360 on a meter because -- for whatever reason, you're going to dose. And if that 360 was really a 240, your child's going to be hypo, we're going to be getting out the gluca gun, and yeah, we

would sacrifice several of the convenience factors. Not that I want to bring back the hanging drop of blood, but yeah, I mean, what's another 20 seconds versus five seconds? Absolutely, accuracy is the most important.

And that's really what I heard across the board. I'll talk about the survey tomorrow, but across the board, parents want accuracy, first and foremost.

DR. HARPER: Ellen, do you believe that parents of Type I diabetic children, or some Type I diabetics that you've talked to, do you believe they realize that the blood glucose meters have some inherent inaccuracy? Or do you -- do you think they're aware of that or not?

MS. ULLMAN: Well, in the survey, and I don't remember the statistics exactly, but it was approximately -- at least 40 percent thought it was 15 percent or less, within 15 percent. And this was a fairly sophisticated group, because they all took it online, so they were all Googling, and they could certainly look up to see that it was 20 percent. So, yeah, I don't think -- I think there are a lot of

people that have no idea, and I do think we need labeling.

DR. CLARKE: Questions?

MS. BOWMAN: My name is Cynthia Bowman. I'm from Long Island Jewish Medical Center. And I was just wondering how you would tie in -- I mean, I realize the total testing process is pre-analytic and analytic and post-analytic. And you know, you have to equate glucose monitoring with wave testing. And the definition is -- you know, the fundamental definition is it's so simple that you won't make a mistake. And if you do make a mistake, it doesn't matter. It was part of the, you know, fundamental definition of wave testing.

But how much do you think that actually contributes? I mean, do you think that still contributes to some sort of devaluing of it? As many fail-safes as you put in, et cetera, do you think that, especially in the hospital setting, that people take it as seriously as they should? Do you think that that actually gets in the way of taking it seriously?

For anybody.

DR. HARPER: You know, I'll leave the clinical comments to the clinicians, but I would like to comment that Carol Benson tomorrow will be giving a talk on sort of the difference between CLIA requirements -- she's not going to go into it a lot, but I'll mention that anything that's cleared over the counter is actually automatically waived. So the evaluation that goes into a device that actually gets evaluated for being simple and having those fail-safe processes isn't actually implemented for all over-the-counter reviews. So that is something that perhaps is not in place necessarily for all over-the- counter product reviews, including -- including this.

Now, that doesn't mean we don't assess any user factors or the ability of users to use these devices. But it definitely is a place where if people are assuming that these devices have gone through that type of evaluation, that actually isn't true.

MS. BOWMAN: I mean, in that grueling hospital environment, one of the things we find is a resistance to proper technique. I mean, these are supposed to be easy. Plus, you know, in their defense,

they're very busy. They're doing many other things, et cetera, et cetera. But you know, doing the QC, paying attention time after time to the technique, et cetera, et cetera, is sometimes devalued. And, you know, I realize it's not just the meter that you're looking at here when you're looking at accuracy. You're looking at the whole process.

DR. SCOTT: I'll take a shot at that, and also maybe try to get a discussion amongst the panel members here, because I don't think there's a single shy person up here, so ....

The total testing process is really more than just the meter. I was talking with Gary Tobin, who heads our diabetes clinics in St. Louis, just last week. And he believes that errors in the meters are just a small component of the total number of errors that are made in the intensive care units. There's errors in dosing; there's errors in timing. But -- and I think Dr. Ginsberg made this point -- the errors of the meters are a small part. But it's sort of like Toyota, you know. They've got a brake problem and they've got an acceleration problem. And to say,

well, we're going to ignore the brake problem, because the acceleration problem is bigger. Fix what we can, I mean one step at a time. And meters are certainly one of those steps that I think we can fix.

I'm tossing that out to Barry.

DR. GINSBERG: Actually, probably the biggest error in glucose monitoring is one that I don't think I had on my slides, although I may have, in the home setting. And probably the biggest single error is hand-washing. That, you know, all the meters say wash your hands before you do a test. Well, it turns out nobody does. We actually studied that, and at least four out of five patients did not wash their hands, including our director of blood glucose monitoring marketing, who had diabetes, and didn't wash his hands before he did it. I came in one day, and our head of a lab was bashing me as I walked in the door, that "your meter is so inaccurate." You know, "I don't have diabetes, and I'm getting a reading of 300." Well, it turns out she had just eaten a banana, and the banana was on her hands. We had to wash her hands, the blood glucose was back down

to 80.

Then if you have Chips A'hoy on your hands, that will raise your blood glucose by 200 to 300 points. That's as bad as talking about with the icodextrins(ph), that hand-washing is a critical factor in these things. There's probably nothing you can do -- nothing you can do in the meter to protect against that. And part of it is that the drop of blood that you take now is so small. For those of you who don't -- I assume you all know this, but if you don't, the average meter is now running .3 to .5 microliters. If you were to look at that on a piece of paper, that's a decimal point. That's the dot at the end of the period is how much blood you need to run one of those things.

Well, it doesn't take very much glucose to raise that drop of blood by a lot, and that's a big error. And I'm not suggesting in the least that we don't need better accuracy of meters. What I'm suggesting is that not everybody needs better accuracy of meters. And a better accuracy of meters means that you have to give up something, that you have to go

back to a hanging drop, or you have to go to a bigger meter, or you have to go to longer times. That's a consumer products issue.

And some people, you'll need to go to hanging drops. And some people who don't like to go to bigger meters, and some people don't care. What's appropriate is to make people understand -- help people to understand what they need and how they get it, and then give them the information they need to make that decision properly.

DR. HARPER: So I have maybe something to add to that, because I agree that certainly user errors and perhaps of unawareness of the things that might impact a glucose result definitely do contribute to error. But what I'd like to point out is that that error is actually not captured in the numbers that we were talking about earlier today in terms of the requirements. So that error is above and beyond the data that FDA, for example, sees when they're clearing or approving a glucose device.

So the studies that we see are performed in the laboratory, or they're performed in cases where

maybe a layuser comes in. And they probably make sure they wash their hands.

DR. GINSBERG: Absolutely.

DR. HARPER: So -- so although we may not be able to address every single source of error, we can certainly try to increase education and awareness. Like Mitch said, we actually -- I'd like to see if we could focus on the types of error that perhaps we can address, because the error that you're referring to is on top of the 20 percent. So what you're seeing is that, you know, you have the 20 percent error inherent in the system, in an ideal situation, plus any added error based on use.

MR. TORJMAN: Marc Torjman, Cooper University Hospital. I think you've probably answered my question. But what I was wondering is whether there are requirements for the manufacturers to actually alert the patient when a glucose value is -- is out of range, and how you define this out of range is an open question, I guess. But does the agency require that, so that a patient who is at home at least has some idea that they need to repeat the measurement, as

opposed to the -- in the hospital setting, where the clinicians tend to make those decisions and send a blood sample to the laboratory to make sure that they have an accurate value? That's not the case with the home use.

DR. HARPER: Yes, that's definitely a challenge. So the home user does not have another method to rely on to check a value, and I think in most cases, a false result may be unidentifiable to the patient.

Probably the only way a patient might be able to determine when a meter is not working, and this is if it's a systemic problem, is if they do the recommended control material testing. We've heard, though, that a lot of patients may not do that as often as perhaps is recommended. So if, you know, a reduced amount of control material testing is done, then they may not catch any inherent problems with their meter. And also if it's a sporadic issue. If it's not hand washing, if it's a sample application issue, if it's something like that that's specific to that one test strip, then I think that there is a

problem that the patients at home wouldn't be able to identify.

DR. GINSBERG: Let me raise a question. We're dealing with two separate items, and we're dealing with them together, which is probably inappropriate.

When we talk about a blood glucose which is really 240 but measures out 360, there we're talking about an outlier. And outliers occur at .1 percent or less, so one in a thousand strips will be there. I mean, they happen, and I think we ought to be paying more attention to them, but they're not that common.

When we talk about a 20 percent error, which we're saying 95 percent of the values would be that, and we do the ISO testing, the other five percent are not out at 400. They're at 21 percent, 22 percent -- they're not way out there. They're pretty close.

On 240, a 20 percent error is 48. So you're looking at 200 to 280. That's a big difference, but in terms of insulin dose, that's only one unit for most people. And that the absorption change in insulin, if you give somebody six units, the

variability of giving to the same patient in the same place with the same syringe or same pen or same pump at exactly the same time of day is 20 to 35 percent. So they're going to be one and a half to two units off just because the insulin's so variable.

MR. TORJMAN: Thank you.

MS. ULLMAN: And I would add that one unit in a two- year-old is huge.

DR. GINSBERG: I'm sorry, I agree with that, because a unit -- she was only taking two units. But in a two- year-old, 40 wouldn't be a unit. Forty would be a quarter of a unit.

MS. ALLBRIGHT: Hi, I'm Ann Allbright. I'm the Director of the Division of Diabetes at the CDC. I'm going to ask the panel a question that I hope will spark some discussion. From the public health perspective, meter accuracy is absolutely critical for patients who are testing for their management, and certainly for hospital use. But there is another end of the spectrum, and that is really the screening arena. And it's very controversial, and those of us in public health deal with this every day.

So comments from the panel about not diagnosis of diabetes, but screening, and where it's used. And there's actually a reasonable amount of evidence that suggests that people are much more responsive for follow- up if their screening conversation involves a blood measure of some sort, as opposed to paper and pencil risk test.

So I'd be interested in the panel's comments on that, particularly as we're now moving into the arena of needing to identify people with pre-diabetes, get them into having conversations, and now developing a national prevention program.

So, eager to hear the panel's comments on screening using monitoring.

MS. BERNHARDT: Well, currently, the FDA does not clear meters for screening or diagnosis. We're well aware that they're used in that manner, especially like at health fairs and stuff, but just informationally, they are not cleared for that use.

DR. CLARKE: If I may say something as a pediatrician who -- whose parents and grandparents -- who has parents and grandparents who are forever

screening their children and their grandchildren, and calling at all hours of the night with blood sugars that are 150 and 160 or 40, and these are in children who don't have diabetes. And this necessitates quite a large workup, because they are either -- their local physician will not see those children. They are sent immediately to the medical center with the diagnosis of diabetes, which they most of the time do not have.

So that I think that if we had something that was more accurate in terms of screening, I think that would be a tremendous asset, but it would need to be in every system, it wouldn't need to -- you know, because it's grandmother's meter that is usually the impetus for the patient to come to the hospital the next day.

DR. GINSBERG: I'm going to disagree again. If you had a more accurate meter and you took it from 20 percent down to 10 percent, what that would mean is for an average value of 120, instead of being 120 plus or minus 24, it would be 120 minus 12. So you'd go from 144 to 132. That's not going to make a big difference in your phone calls.

And as to Ann's question, I think that's -- it's a very interesting question, that when you look at impairing glucose tolerance, which I think is one of the major risk factors that you're looking for, an hour after a meal, you're looking for a blood glucose value of 200 or more. And so if you take the 20 percent accuracy where we are now, that would say that if you're looking for 200, your 95 percent confidence limit is 160 to 240. The upper side is no problem, you still have over 200, you still have impaired glucose tolerance.

The problem is, what happens if you're low? Well, I would actually say, you know, if somebody two hours after a meal has a glucose value of 160, they're still somebody you want to look at. So I think while they're not as accurate as you'd like for this, I think they still meet the needs.

DR. KLONOFF: I'd like to comment on this discussion and where I see this is going. I feel that the group has already accomplished a lot. When this meeting was announced about three or four months ago, my sense was that there were many people who thought

this meeting could be very confrontational and that perhaps the FDA would be making demands of industry that it would be very difficult to achieve. Perhaps industry would be refusing to budge on the products that they're producing.

In fact, I'm seeing this as very collegial, and for the most part, it seems as if the two groups are coming together very well. Plus the medical and academic communities, everybody seems to agree that we need in general lower standards for accuracy, go down from 20 to around 15. That's the number that keeps coming up.

And what I think is happening is that when the FDA is setting standards, these are regulatory standards. A regulatory standard means this is something that's achievable. When doctors make requests or standards, we tend to talk about what's needed. Those are clinical standards. We can say we're like the ADA, we want five percent accuracy or we want zero percent errors, we want absolutely perfect meters. We can say this, because this is what we want in our hearts. But that's not necessarily

achievable, either in terms of the technology or in terms of making it economically viable. I mean, for enough money, you probably could get a device that's extremely accurate, but the cost would be impossible.

So what I think we're seeing here is a very healthy process. I see that industry is responding to the -- the sense that they're hearing from FDA, from what they're hearing from the clinical and the medical community, that something has to be done. The standards have to become at a lower number. And there are some issues that still have to be resolved, such as do hospital use meters need the same standards as outpatient use meters? That's probably going to be discussed some here.

The kind of process that we're seeing reminds me of how standards are set in other industries. I'll give you an example. I'm from California, and almost every hospital in California now is being replaced because of the seismic act of earthquakes. And we realized in California in the '70s after the San Fernando earthquake that we needed some hospital seismic security activities, and the

Hospital Seismic Security Act was passed right around the time that blood glucose monitoring came to the forefront. And as blood glucose monitor standards have been getting tighter over the years, so has hospital seismic security.

And what happens is that about every ten to 20 years, it's announced that every hospital is going to have to withstand an even greater earthquake, or they have to shut down. So essentially, all the hospitals get rebuilt. We've just seen the most expensive hospital in the history of California built, which is the Ronald Reagan UCLA Medical Center. That cost over a billion dollars. What'll happen in another 20 years is the earthquake requirements are going to get even stiffer, and even modern hospitals like Ronald Reagan, my own hospital, Mills-Peninsula, and many others are going to be deemed behind the times.

This is just a natural evolution as the technology improves, and it's going to happen with blood glucose monitoring. But what I really have enjoyed about the meeting so far is seeing that people

really are getting together. We see the need. We're not going to get to zero percent today, just like we're not going to get to a medical center in California that will withstand every possible earthquake. But it's a gradual process, and every once in a while things get tighter and tighter. That's what's happening with blood glucose monitoring.

So I wanted to really commend what I think must e a lot of really good behind-the-scenes work by FDA people and by industry people and AdvaMed to get to this collegial point. Because it didn't automatically seem like it was going to happen, but it is happening. So, thanks to everybody.

MS. RUTHERFORD: Diane Rutherford again with Ken Block Consulting in Dallas.

What I brought up earlier was the dose accuracy issue with syringes. What you had said the difference between one unit and two on a child, is very significant. But if I recall correctly, the tolerance on one unit could -- you could be giving two and still meet the syringe requirements. So I would think that would still be a concern on some level.

Also, you guys are talking about re-testing high values, it seems to be. And basically, it sounds like if you get a value you like, you accept it. If you get a value you don't like, you're going to re-test and see if it's really the right number. So how many of the numbers that we like are actually accurate?

DR. CLARKE: That's an exceedingly important question, and who would like to answer that?

MR. CEMBROWSKI: Hi. George Cembrowski, University of Alberta. The test is as good as the drop of blood that is derived from the patient. I'm wondering if someone has done any theoretic calculations as to how well we can measure this drop of blood as the volumes get smaller and smaller and the analytic time decreases, as well. I think we're hitting a wall, and I'm thinking that there are all kinds of reasons for bad results. A hypertensive patient in Alberta winters, the drop of blood is mercilessly evaporating once it comes out of the patient; the nurse might not be all that good at moving that drop of blood to the instrument. I'm

wondering if, especially in the ICU, whether we can even achieve 100 percent of the time, a CV of five percent measuring a drop of blood. And I'm wondering -- Dr. Ginsberg is probably the numbers guy. Could you throw some numbers around?

DR. GINSBERG: I actually think that a drop of blood is probably a bad way to measure blood in an ICU. That you have patients who may be hypotensive; they may be over-hydrated, they may be under-hydrated, and all of those add a lot to the inaccuracy of a drop of blood in ICU. I think a venous and arterial is probably a much better way to do it in ICU.

Unfortunately, not all meters are calibrated appropriately for venous/arterial blood. As we've talked about today, a number of the meters are very oxygen- sensitive. All the glucose oxidase meters, all the oxidase biocenters are, or should be, oxygen-sensitive. If the oxygen is high, the reading is going to be low. If the oxygen is low, the reading is high. They are calibrated to capillary blood, but capillary blood is a bad way -- I think a bad way to go in the ICU.

That said, I'm not a hospital expert. I haven't done anything in the hospital for over 20 years.

MS. HARPER: So, if there -- if there were a need for, and it sounds like there is, obviously, because these products have migrated into the ICU and other hospital settings even though they're not -- they're tested or intended for those facilities. Since there is a need there, what type of -- you know, we're talking a lot about requirements. We've heard a lot about, you know, that maybe ten percent accuracy might be a minimum needed for that type of environment. But how do you design a study -- I'd like to hear from some of the -- the clinical people - - how do you design a study to demonstrate that in that population in the ICU, that you have a device that performs adequately?

DR. SCOTT: They are difficult studies to do. I mean, your outcomes are going to be fairly rare events. I mean, look at NICE-SUGER. They had a three percent difference in mortality, and it took 6,000 patients to determine that. But, I mean, ideally what

you would do is randomize patients in units to current meters versus something that is far more accurate, and track outcomes.

But I think the end required for that study is going to be quite large.

DR. HARPER: So that's definitely true for clinical determinations of benefit or perhaps accuracy. But I'm also interested in hearing how you believe that, you know, because I agree that perhaps capillary blood might not be appropriate in some very sick patients, or things like that. Analytically, how would you determine for the range of types of patients that might be seen in hospitals just what type of analytical study? Because right now we do, as Pat pointed out, you know, we have 100 patients, and they do 100 capillary samples, and that's the performance. Is that enough for something in the ICU, or not? That's ....

DR. SCOTT: I think the closest that addressed that was a study that Brad Carrone (ph) did at Mayo, where they simultaneously drew all three types of samples and then sent a venous sample to the

main laboratory. And I presented those findings, or actually the capillary, which surprised me, because I tend to agree with what -- who was it, was it George that said that finger-sticks may not be the ideal sample in intensive care units. But that was a reasonably well designed study, in my opinion. And I think it's going to be very meter-dependent. That was done with one of the newer meters.

DR. KLONOFF: One feature of a study that was suggested earlier is that you make sure that there are enough data points in the hypoglycemic range. That's what we did when we had our continuous glucose monitoring performance guidelines through CLSI. We ended up stipulating that there has to be some hypoglycemic points. Otherwise, you get this, you know, 180 to 240 syndrome, and you can certainly look just fine on the error grid, and you don't even look too bad on the analytic. But you have to have some hypoglycemic points.

DR. GINSBERG: The other point in terms of ICU patients is, the hundred patients that you do for a supplementary blood glucose study is because the

assumption, with some reason behind it, is that out in -- at home, that most patients are fairly healthy and are similar, although you do ask that you have a wide variety of races involved.

When you go to the ICU, I think a hundred patients in insufficient. I think you need to start categorizing by kind of patient, and look at the various kinds of patients you're going to have in terms of hypotensive, under-hydrated, over-hydrated, and so on, as well as a wide variety of blood glucose .... So I'd be surprised if a hundred was enough.

MR. MELKER: Richard Melker, University of Florida College of Medicine. I could change the tone of the meeting, but -- (laughter from audience) -- I've elected not to. But number one, the first thing I'd like to say is that everybody talks about meter accuracy. It's really the test strips that are the issue. It's not the meters. The meters are very accurate. It's the test strips that deteriorate over time. So the first comment I'd like to make is, nobody's mentioned glucose control solutions, which, if you go into the MARE databases, the reason that all

the manufacturers use when a patient gets an inaccurate reading and gives themselves an inappropriate dose of insulin. I -- I've done a lot of studies on myself, since I'm a Type I diabetic, and so I have to consent myself for studies, which is fun. But I'll actually run my glucose up and down on purpose in order to test glucose meters. And so I'd like to ask Dr. Ginsberg, if I open a vial of test strips and I test the first time and the number is 175, and I immediately take another drop of blood and put another strip in the same meter and it's 200, and then I immediately take out another strip and I put in the meter, it's 225 -- if I had taken any one of those individually and tried to calculate that it was plus or minus 20 percent, you realize how confusing it becomes to the patient when they only did one of those three. I'd like to say that the middle one was the average of the three, but you don't know that if you only take one, and one of them is 175 and one is 225.

The last thing I want to say is what you said about hand washing, which I think is really interesting. Because if you wash your hands and you

don't dry them really well, you get low glucose readings, because you have water on your hands. So not washing your hands is a problem, and washing your hands and not drying -- which takes a fair amount of time to do properly -- is a problem.

The other problem with not drying your hands completely is if you open the vial and you take a glucose test strip out while your hands are wet, you can ruin all the other test strips in that vial. Nobody teaches patients about any of these issues, so have at any of them.

MS. HARPER: Well, I personally really appreciate those comments, because these are things that, you know - - we always struggle with labeling. We heard some really good comments today that may help us with that, including this one. Labeling isn't, obviously, the -- the be-all, end-all, and it certainly isn't the answer to all. But the more input we can get on how we can help to develop clearer labeling for patients and educational materials, the better, because this is certainly something that adds to the problem.

DR. CLARKE: And take it from someone who one night was testing a continuous glucose sensor, and it was -- the alarm went off to put in the glucose reading, and we were just finishing up dessert. And it was this wonderful English, what is it, you know, trifle, that's right. And I licked my finger, and got a glucose of 222 and nearly passed out, and wondered what in the world I was going to do next. Yeah, washing your hands is really an important thing to do.

DR. CARISKI: Alan Cariski, LifeScan OPS California. I think this is an issue that's been addressed tangentially, but I just wanted to make it a little clearer and get some comment from the panel. It's true that the accuracy standards are plus/minus 20 percent, plus/minus 15, but that's for populations, whereas the precision is closer to five or six percent, so that the variability that any individual patient will see is generally going to be a lot less than plus/minus 20 percent, because the things that affect the accuracy of the strip are going to be pretty constant to that patient -- the hematocrit, the uric acid, et cetera, et cetera. I was wondering if

the panel agrees or disagrees. Thank you.

DR. HARPER: Well, for some of that, it depends on if you're comparing two results or not. I mean, if the patient has a hematocrit, it may actually affect the true value. I mean, it may actually affect the reading.

MR. MELKER: (Inaudible, off-mike) the next glucose is going to be off by the same amount?

DR. HARPER: Right. But your comment that -- you're commenting that this may not be a problem with 20 percent total error in an individual patient, and I'm saying it's possible it could be that or more if they have a high or low hematocrit. So it could actually -- or an interfering substance, or something like that. So, it's -- imprecision is certainly part of it, and certainly where there are not other issues, it is part of the total error. But it isn't necessarily the only part of the total error that might lead somebody to treat on a number that isn't really the true number.

DR. CLARKE: Next?

MAJOR MANN: Major Mann again from the Army

Burn Center. One thing that really concerned us when we discovered the hematocrit effect on our point of care glucometer was when we approached the lab, and we use CLIA -- or we use CLIP, rather than CLIA, actually, in the military -- they said, well, the meters meet all of our standards. And we looked at the fine print, and they were supposed to test the meters on normal subjects. And our subjects, our burn patients, actually have a hematocrit of 24 percent. And so that is below the meter accuracy that's on the label of our glucometer. So there's a big disconnect in the providers understanding the variability within these meters, as well as the meters that we tested that said that they have an accuracy to 20 percent hematocrit. Frankly, that wasn't true, either, in our data that we collected.

So I guess one of my concerns would be that there would be a requirement within the hospital, within an ICU, to test the device that you're going to use on the patients you intend to use it, not on normal controls.

And furthermore, as I had mentioned before,

I think a study that you can do -- an outcome study clearly is not feasible to do any kind of morbidity or mortality. But it's very simple to do an outcome study on rates of hypoglycemia using different meters. And that's exactly what we did, and we found that once we changed to the corrected value, we stopped having occult hypoglycemia, and this means the matched blood that we sent to the lab was hypoglycemic, but the point of care device was clearly in a normal range.

So I think that's a very easy study to do to test meter performance in a variety of subjects.

Thank you.

DR. GINSBERG: Let me actually comment on that.

Hematocrit is actually much more complicated than that. Part of it is that no meter is a pure whole blood glucose meter, and no meter is a pure plasma meter. They are all somewhere in between. And so based upon an average hematocrit, they then put in a correction factor to bring you to a plasma value. And if the hematocrit's not correct, then that correction factor needs to be corrected. Many meters

have a correction factor within that, if they can measure the hematocrit.

But it also turns out that hematocrit -- that the change in the volume of red blood cells can also affect the reading in other ways. For example, the surface area of the electrode is a critical factor in terms of current flow, which is the determinant of -- of the blood glucose. The higher number of red cells you have, the more red cells clog up, because the red cells don't pass that current -- clog up that electrode, and you have to correct for that as well.

And for at least some companies I've seen, the error in hematocrit can be huge. Not four or - when you correct plasma to glucose, that's a 12 percent correction at a hematocrit of 40. The correction can be up to 50 percent, depending on the particular meter, that meter/strip combination that you're using.

So I applaud that you're correcting for the correction from whole blood to plasma and changing the hematocrit in that, but the correction is actually more - - and you're doing something, but the

correction is actually more complicated than that.

MAJOR MANN: If you're looking at a similar patient population that all tend to be anemic, wouldn't the normal correction sort of be consistent across that population?

DR. GINSBERG: Yes.

MAJOR MANN: You wouldn't use the same correction factor for a neonatal population as you would for an anemic adult population --

DR. GINSBERG: Right.

MAJOR MANN: -- for example, so that you are precise and maybe not accurate, but --

DR. GINSBERG: Yeah. I -- I'm sorry. If you were correcting using empirical data, I agree. That's even better.

MAJOR MANN: Right.

DR. GINSBERG: I assumed you were correcting just based upon taking the 12 percent for hematocrits.

MAJOR MANN: We -- we actually take the daily hematocrit. It is automatically extracted from our patient care record in the electronic medical database, and the formula is embedded in our

electronic medical record. So when the nurse enters the glucometer value, it automatically retrieves the most recent hematocrit value and matches to that most updated value. And we feel like a 24-hour or 12-hour hematocrit is pretty fine in a non-bleeding patient. And we again have had a decrease in the rate of hypoglycemia that's measurable compared to a unit that didn't correct, and their rate of hypoglycemia actually went up.

So that's just a different type of study to think about.

DR. KLONOFF: I'd like to comment.

First, I think what you're doing does make sense in a specific group of patients. Second is we need to see more modeling data like what Marc did with Boris Kobichev. If we're going to find out what are the -- what are the clinical consequences of error, then either we do a study which basically, as Dr. Scott said, is impossible -- it's unethical, we're never going to do that -- or we get some modeling data.

Modeling data is so useful. You can find

out what's the frequency of hypoglycemia, what's the frequency of different combinations, complications. The modeling studies do require assumptions. Even with all the equations, they're assumptions. And I think that it would be very nice if others in this room were to try to do a study similar to what Marc and Boris did. Their study is the first of its kind that's gone beyond looking at what happens to the insulin dose with inaccuracy, but actually what then goes on to happen to the blood glucose level. So it's a pioneering, very important study.

But the basic idea of modeling what happens, others could do it, too. And I think that if we see more of this type of information brought forward, we'll all have a better idea. Because we're saying, well, we'd like more accuracy, but I think it's still very difficult for all of us to get a handle on how much accuracy do we really need, let alone what's it going to cost. So let's see some more modeling studies.

MR. SOUTHERLAND: Phil Southerland, Team Type I. I've got a few comments which I hope will

maybe spark also some discussion.

I've had Type I since I was seven months old, so almost 28 years now. So Ellen, I can relate to what your son's going through. But in the first 25 years, I checked my blood sugar about 118,000 times, which means at .01 percent of outliers, I've had 118 outliers, and then guess what? I've lived through them.

But, you know, we've come a long way in that time. My mom used to squeeze the urine out of my diaper to get it onto a test strip to find out where my blood sugar was four hours before, you know, where I actually was. So here we are talking about a small differential of percentages, but I do agree, David, to what you said. We can keep moving the needle forward, which is going to make the control better.

Also in the last three years, and using continuous glucose modeling technology, I've checked my blood sugar 109,000 times. Averaging the fact that I pressed the button about a hundred times a day to figure out not only where I am, but the trend in where it's going. And if you look at the average person who

checks two times a day, going on the basic math, that's .04 percent of the day that they know where the blood sugar is. You know, they have no idea what the trend's going, what direction it's going, and that creates a huge margin for user error and the effect that it might have.

So can we, you know, start pushing for the Type I community who's on insulin using CGM, because that takes so much of the user out of it the majority of the time? And I do agree, there is still user error. Last Friday in advance of coming to this meeting, I checked with three different meters. One of them said 69; one of them said 150; another said 213, all at the same time. So I checked again. Then another one said 78; one said 108; one said 113. And the third time, they were all around 102. You know, it's --

UNIDENTIFIED: Wash your hands?

MR. SOUTHERLAND: I did wash my hands before that third test.

(Laughter) But, you know, that user error, if I went on that first 213 and gave three units of

insulin like I normally would, I'd have been in for big trouble if I wouldn't have checked my blood sugar 45 minutes later like I always do after I give insulin. So, I mean, where is -- where's the curve, where do we need to go? Can we make this, you know, five percent a little bit better?

The ICU, I've been in the hospital, and they said, we're going to check your blood sugar every six hours, and we're going to give you a shot of regular insulin every six hours. You know, we're talking about five percent in the meters, when I'm going to use regular insulin, which I haven't used in ten years? Where's the margin of error in that one?

DR. HARPER: Yes.

MR. SOUTHERLAND: Where do we draw the lines?

DR. HARPER: Those are all good comments, and I'd like to say, we are talking to -- not only do we work with the glucose meter manufacturers, but we also work with the CGM manufacturer community to try and get them accurate enough so that people could possibly in the future use them.

But I think you raise a good point about standardization, and that came up earlier in one of the talks. So I would like to hear perhaps from some of the other panel members on whether there's a need for increased standardization or not, given that at least at home, single meters are used, and generally multiple different types of meters are not used. Or are they?

DR. KLONOFF: Phil, the problem that I think happened to you that day was your 213 was one of these five percent outliers that was really bad. And this seems to be an area that people are now starting to discuss. It's only in the last year or so that I'm hearing much talk about it. And if -- I think something should be done to address that five percent group. And it has to either be -- at the most extreme would be to eliminate it completely and say 100 percent have to be within range, or something that may be more achievable is to just analyze that five percent and say how it has to be distributed. Currently, the worst of that five percent can be as bad as you can imagine. But if something was done to

tighten those up, those you would just never see a number like that again.

MR. SOUTHERLAND: And thank you, everyone in the room, for all the work you're doing. You're making all of our lives better. Appreciate it.

DR. BRETON: There was one -- one more thing that I wanted to add, in that Phil is pointing at an application of SMBG that we've barely discussed. And they're used to calibrate the continuous meters. And all right, you know, an error of ten or 15 percent might not make a huge difference in your insulin dose, especially if you have error in what's absorbed or what's actually injected.

But an error of ten percent will have a dramatic effect on the calibration algorithm of a continuous meter. And that effect won't be apparent when you calibrate; it will be apparent at your next meal excursion. And now you're not talking about ten percent error any more. You're probably talking about 50 to 100 percent error. And if you ever were to, by all means use a CGM to treat or even take another SMBG at that time and that happened to be erroneous, too,

you can be in actually serious trouble.

So there is also clinical consequences of SMBGs that are not directly linked to your insulin dosing at the time of SMBG.

DR. GINSBERG: I'm going to take that, actually, as another chance to suggest what I'm suggesting, and that is, there are different segments. For that segment that is looking to calibrate their glucose meter, to calibrate their continuous monitor, they may want five percent inaccuracy, or five percent/95 percent confidence limits on inaccuracy. That's not the suggest that everybody has to have that. And a suggestion that there are different needs for the device, and that we ought to publish what the accuracy is people can select the meter appropriate for -- system appropriate for their needs.

DR. HARPER: But part of the -- part of the thing we struggle with there is how do we get data good enough to actually demonstrate the accuracy? Because I think what we struggle with not only in blood glucose meters, but in any of the devices we

regulate, is that we get the data we get. That's what goes in the label. And in the field, we may or may not see that same performance. So what -- what we would like to hear from people, either today or, you know, you can send in comments or give us a call or whatever, what we actually want to know is if you have ideas on how we can actually tell people how they work, for real, in the real use environments.

MR. BALTUR(ph): I'd like to make a comment about ICUs and the intensive insulin therapy. The nurse who runs this usually makes the decision based on the current evaluation and the one that would probably make six hours or two hours or whatever the interval was. Some of us think that a continuous glucose monitor will be ideal for intensive IV insulin therapy in a ICU, provided you had a reliable machine. And we tested the continuous glucose monitor which is interstitial, and that's not good enough, because the delay is so bad. And I know there are some companies, and I tried to get a hold of them, that have a meter that you can place in the vein, and then you really have -- there is no nursing, there is a continuous

monitor; you can look at its trend, and you can make very nice decision.

Can you tell us what the status of those kind of instruments are today?

DR. HARPER: There are some blood gas analyzers there are indwelling currently, and I believe there is one system that was, I believe, cleared in the '90s that is capable of making frequent measurements of blood glucose meters. But in terms of continuous monitoring, I think, you know, we have also heard about some companies that have some products in development, but none of those have been recently cleared.

MS. PINKOS: Arleen Pinkos from Baltimore. We have heard 15 percent over and over again, and there does seem to be some kind of agreement or comfort level with that. So my question is a couple. Is 15 percent good enough for -- I would like to hear from all of -- at least the clinicians on whether 15 percent is good enough. And also not just the 15 percent, but how many milligrams per deciliter should it be below 75 milligrams per deciliter?

And after you've answered that, I'd just be curious to know how you would apply those to home settings versus clinical settings. Are they the same? Should they be different? But is 15 percent good enough? What about the low range, too?

DR. GINSBERG: I think it depends upon who you are. I think if you're calibrating a continuous glucose monitor, 15 percent is not good enough. I think if you're an intensive insulin therapy patient, keeping your blood glucose in the 80, 85 range most of the time, I think 15 percent is not good enough. I think if you're a Type II on insulin, sure, 15 percent's plenty good. If you're a Type II not on insulin, I think you don't even need 15 percent, so it's certainly good enough for them.

I think a reasonable number below 75 is to use the same standard that you use for 20 percent -- that basically you take the value at 75 and just bring it down. And at 15 percent, the value at 75 is 12.25 or 12, and you just bring that down from there.

DR. CLARKE: Let me say that I have a little problem with that 15 under 75, and I think that if --

I think that 10 under 75 would be much more acceptable, especially in children. And -- yeah, and I also think that, now, I agree with your other comments. I think that if you're going to put labeling on glucose strips, that it really needs to say specifically, under 75 this product will read this. Above 75, it will read that. Because those are the standards -- those are the ISO standards that you're using, and to just tell somebody that it's accurate within -- this many times within this percent of accuracy really doesn't tell them anything that they can use.

MR. BRETON: I would like to add one more thing towards that Bill was saying, which is that for hypoglycemia detection, so for values in the low range, 15 percent still gives you the opportunity to miss hypoglycemic events in quite large numbers. Ten percent seemed that you were actually starting to reduce the number of mis-events, and in quite a dramatic way.

Now, if you're at 150 milligrams per deciliter, the difference between ten and 15 percent

is probably not that great.

Another comment that I wanted to make is that I'm wondering where that fixed rate of error below 75 milligrams per deciliter came from. Because it seems to me that even though it's technically challenging, I understand that, I want to be much more precise. At 50, even, I want to be at 75. And the relative nature of the error that we have higher than 75 seems to still have some relevance below. And so I'm pretty sure that people might have comments about that.

DR. CLARKE: Anybody else on the panel want to take that?

MS. PINKOS: Can I just have one follow-up for Barry? You said 15 percent is probably not good enough for those patient populations. What is?

DR. GINSBERG: Oh, for that patient population, Marc and I are actually talking about the same problem. And that is the problem of knowing when you're hypoglycemic. If you remember his graph of how often you would miss hypoglycemia at various error levels, what he said is at five percent you won't miss

it at all. At ten percent, you'd miss it at one percent. At 15 percent, you'd miss it at five percent, and at 20 percent, I think you'd miss it at ten percent.

And so the big cutoff there, it'd be nice to get down to five percent. That's really quite challenging. I think ten percent would be acceptable.

DR. CLARKE: Neil?

MR. WHITE: Neil White, St. Louis. And I -- I'll say in many ways, I agree with Barry. That's probably pretty unusual. But -- (laughter) -- we need -- we probably need different standards for different things. But I just want to point out -- maybe this is repetitious of what others have alluded to -- if we're ever going to move in the direction of a closed loop, and if we're going to count on blood glucose meters as the standard by which we calibrate the sensor, we can't tolerate large levels of error on both devices in a device that's going to be making decisions independent of our own human input.

DR. BRETON: To further that comment, so we -- at UVa, we have a few trials going on about

artificial pancreas and closed loop. And what we've noticed is that a CGM calibrated with a YSI, for example, is tremendously more accurate than a CGM that's calibrated even with a good meter value. And by -- I mean, you're always have the problem of the delay, of course, and it's still slightly different in plasma or glucose, what you're measuring. But a perfect calibration makes all the difference with the use of a CGM in a closed loop system. And so I really want to emphasize the need for accuracy when we talk about calibration, and not necessarily insulin dosing.

DR. HARPER: So I have a question, actually, for Steve and Barry, so our industry panel.

DR. GINSBERG: May I?

DR. HARPER: Sure.

DR. GINSBERG: It's critical to realize that the YSI is not a perfect instrument. The YSI actually has about a two and a half percent error, meaning the 95 percent confidence limit on that is plus or minus eight percent. So that when we talk about a calibration device, what we're saying is we need something between five and ten percent, which I think

everyone would agree with. And we're not that far. If you look at the very best meters out there, they're now getting 75 percent of their values within five percent.

DR. BRETON: All right. I might want to just qualify a little bit what I said. When I said we calibrated with YSI, what we do is we have duplicates, and actually four total membranes. So that eight percent is dramatically reduced with that type of setting.

DR. HARPER: So, I have a question for our industry panel members, so Steve and Barry. We talk a lot about sort of, oh, what would the ideal, you know, level be. We hear a lot about 15; now we're starting to hear about in the hypoglycemic ranges, getting down to ten or even ten percent, I heard.

From an industry point of view, we'd like to hear some discussion about the barriers to achieving better accuracy. So what would need to happen, from a manufacturing point of view, or a development point of view, to enable meters to become more reliably accurate to the point where we're discussing?

DR. BROTMAN: That's obviously a very important question. And, you know, it's -- manufacturers really look at designing these things and trying to meet the users' needs all the time. And, you know, you reach a certain level sometimes where you have tradeoffs between accuracy. Accuracy is made up of accurate results sometimes, and you're also looking at user -- end user errors and interferences, and so forth.

So it's a delicate balance between that situation. And if you're looking for consistency across the board, as you, you know, as you get closer and closer to ten percent, there may be devices now that are measuring ten percent. There may be -- as you get closer and closer to that non-tolerant level of having any errors, you're looking at, you know, how much these improvements can be sustained across the board. And I think that's a hard thing to do at a certain point. So you get down and you get down, and --

DR. HARPER: Yeah. So what's hard about it? I think we're trying to understand what actually needs

to happen. Are there -- are there any ways that scientists can help --

DR. BROTMAN: Well, I --

DR. HARPER: -- or any ways that we can actually help?

DR. BROTMAN: Right. I mean, I think you're also looking at all the other things that have been brought up today. You have a lot of interferences, and you have a lot of use type of situations. So your, you know, your hematocrit, your humidity, everything else that -- that's in there contributes to this process also.

DR. KLONOFF: Courtney, one thing, it strikes me that we've been hearing about the tradeoffs, that there are many nice features of monitors, so it's not -- even if the accuracy hasn't improved that much, there's a lot of nice features. And some of them, I think, could be sacrificed if it would lead to more accurate blood glucose readings.

For example, in the hospital, you're not worried about, does the consumer like the product? Does he find it easy to use? Because it's not even

the consumer, it's the nurse, and the consumer doesn't really notice what's going on. Or the same thing with a long-term facility.

So if a company is going to have a second line of meters, let's say the standard accuracy and the greater accuracy, to me the first thing they would do is get rid of some of the features that are nice for being less accurate, but aren't necessary. For example, in some cases, have a longer measuring time; in some cases, have a larger drop of blood; in some cases, require a type of coding system -- require it, which is very clunky and difficult to do, but the patient doesn't have to worry about it, the nurse is going to do it.

So I think that some of the ability to create a more accurate monitor has to already be there if certain populations are willing to forego those nice features. And I think those populations will. I'll be even the CGM user, knowing how important it is to have an accurate calibrating device, would be willing to give up some of those features, too. And the majority of patients would never have to face

that. They could get by with the standard accuracy.

DR. GINSBERG: Let me talk about the company I know best, which uses something called dynamic electrochemistry. And what that means is the following. What they do first is they bring -- what a normal meter does, is it brings the voltage up to, let's say, .3 volts, and then measures the current over time as a way of measuring blood glucose.

What this system does is it brings the voltage up to .01 volts, or some number like that. And from that, you measure nothing except the well size. So if there are any changes in the manufacturing of the well, you immediately pick that up.

Then it goes up a little higher and does some tricks to pick up the hematocrit. Then it notices the slope of the current as it's changing its voltage, from which it can pick up the oxygen concentration and the altitude.

Then it goes up to a voltage just a little bit below, where it'll pick up glucose, and picks up all the interference. Then it goes up and picks up the glucose and uses software to get rid of all the

others. And by doing this, they took an original strip, a Korean strip, which had by itself an inaccuracy of 11 percent, and brought that down to a little over five. And they believe, at least, when they designed their strip themselves, that they can get their inaccuracy down to about two.

DR. CLARKE: We have time for one comment. Go ahead.

MR. COMBS: Art Combs from St. Louis. A few years ago, they closed the major east-west road in St. Louis, and so that road right next to the university got a lot of extra traffic. And it always had a 30 mile an hour speed limit, and nobody ever paid attention to it. But now with the extra volume, they put up a new sign, and the new sign said, "30 miles an hour no tolerance." And they started giving out tickets for people driving 34 in a 30. Now people drive 30 miles an hour on the street. I think if you walk into a classroom and say, Passing is 65, but starting next week, passing's going to be 75, all the same people are going to fail, and now more people are going to fail.

My suggestion would not be to lower the standard. My suggestion would be to enforce the standard you have, and enforce it in the user environment. I know that any of the large strips companies can perform a clinical trial, and under the right conditions hit that Clarke A space 95 percent of the time. But that's not what's going on in people's kitchens.

So why don't we enforce the standard, rather than change it? I think that speaks to the outlier question, particularly.

The second thing I wanted to say is, no monitor, in my view, is worth anything unless it's a safety monitor. If your blood sugar is dangerously low, that's an emergency, and everybody needs to be able to detect that with a certain accuracy. If your blood sugar is 300 or 360, it's not an issue from a treatment point of view, and no one's life is in danger. The 20 percent standard is adequate.

So we seem to be in this meeting looking for a one size fits all. We need a standard across the entire dynamic range of glucose, and we need a

standard that's from the entire dynamic range of patient experience from their own kitchen to being profoundly ill in an intensive care unit, and I don't think that's the case.

The last point I wanted to make: I think that from an FDA point of view, and I certainly can't speak for them, but two things have been lost. One is, what is the intended use? And the second is, what is the indication for use? There are many diabetics, and I suspect the young man who was up here earlier wears an insulin pump. This is a $5,000 device. It's very precise. It's programmable, it's complex. Not every diabetic gets it. Some people get a $5,000 device, some people get a bag of 100 syringes for $9.99 and a multi-use vial of insulin. The same thing should be true of meters. That's why I don't believe you can use the same thing in the intensive care unit as you can in your kitchen.

But why don't we view standards, instead of saying, Gee, it's 20, should we lower it to 15? Let's start with, Why don't we enforce the one we have now so we don't have outlier problems? Why don't we make

sure it's always a safety monitor, at the very least, in the hypoglycemic range? And why don't we start thinking about indications for use? If I'm an ambulatory or otherwise well Type II diabetic in my kitchen, do I need an ICU monitor? And should I have the same standard?

Thank you.

DR. CLARKE: Thank you for your comment.

We're going to have to stop now. I want to thank all of the panel members for your participation, and the audience for your participation, as well.

(Applause)

DR. HARPER: So first, I'd like to extend warm thanks to Dr. Clarke, who did a wonderful job moderating Session Number 1. I think we had a lot of good discussion, and I know I can certainly speak for myself that I heard a lot of interesting insights that I think will be very helpful to us.

So while the Session 1 panelists are coming to their seats, I'd like to introduce the moderator for Session 2. Dr. Gary Myers is the Chief of the Clinical Chemistry Branch in the Division of

Laboratory Sciences at the Center for Disease Control and Prevention. In his more than 30-year career at CDC, he has focused on improving lab measurements of biochemical markers used to assess chronic disease risk, with particular emphasis on cardiovascular disease and diabetes.

Dr. Myers is a member of the American Diabetes Association's Insulin Standardization Working Group, and in 2007, Dr. Myers served as president of the American Association for Clinical Chemistry. He has authored or co-authored more than 80 peer reviewed publications and chapters.

Please join me in welcoming Dr. Myers.

(Applause)

DR. MYERS: Thank you, Courtney.

Okay. It was a pleasure to be here, and I want to thank Courtney and the FDA for asking me to moderate this session.

Our session today is, the second session is Blood Glucose Meter Performance, Interferences, and Limitations. And you've heard from Dr. Scott and Dr. Ginsberg when they were talking about total error, the

idea that total error is not just bias and imprecision, but there are other things as far as the pre-analytical and post-analytical that goes into considerations of performance for blood glucose meters.

So meter performance is limited by a variety of environmental, physiologic, and operator factors. And in my brief introduction, I'm going to list a few of these things. You've heard some of them already described. We're going to refocus on these in this session, so our speakers will have an opportunity to go in a little bit more detail.

What I've listed here, and I'm not going to go into any kind of detail, but just list these -- there are possible factors, interferences, that have been investigated and reported in the published literature. And there are environmental factors. We've heard about exposure of test strips -- exposure to air, exposure to light. The use of generic strips -- that may not have been evaluated on different meters, but are also available. Age of strips -- we did an evaluation of meter performance at the CDC some years

ago, and we found that on getting strips from the pharmacy, if you had strips that were early in their shelf life versus those that are towards of their shelf life, we found a significant difference in the performance of those strips. Reuse of strips -- altitude has an effect on the performance; humidity and temperature, again, having the strips exposed to various humidity and temperatures. Specimen preservatives -- volume of the sample, are you under-dosed or over-dosing the sample? And animal blood -- not in the measurement of animal blood, but in the use of animal blood in the preparation of control and survey materials. Maybe not as prevalent now as it used to be, but there was a time that animal blood was used to prepare control materials and survey materials. That may still be going on. And we know, we've heard from Dr. Scott, the problem of matrix effects when survey materials do not simulate patient samples, so that's a big concern when evaluating the performance of meters.

Patient-specific variables: extremes in hydration. Is the person hydrated or over-hydrated?

Hypotension, prandial state. We've heard a lot about hematocrit, and we're going to hear some more from one of our speakers. Hemodialysis, hemolosis, extremes in pH. Cases of severe acidosis, like diabetic ketoacidosis, as an example. Hypoglycemia -- you've heard a lot about hypoglycemia from Dr. Sacks. Blood sources -- differences in arterial versus capillary, versus venous blood. The specimen matrix -- differences between whether you're measuring it as plasma or whole blood.

And then the issue of substances and conditions.

Acetaminophen, which at therapeutic drug levels certainly causes interferences. Ascorbic acid, dopamine, fluorescein IV's, mannitol, salicylate. We all know about the therapeutic products with non-glucose sugars:

maltase, galactase, and xylose with GDH and PTQ test strips. And we're going to hear a little bit more about that in some detail in just a minute from Dr. Harper.

So as we go forward in this session, just

some issues and questions to consider. And some of these we've already heard discussed and brought up in the previous session. But again, focusing on inferences and limitations, some things to consider.

First of all, the investigation of interference effects is the responsibility of instrument manufacturers. This really has to be done by the manufacturers before it's put on the market.

The information provided is often too vague and to be of little value, and some manufacturers may take exception with this, but this is some of the general feelings that have been published in the literature by people that have done studies: There is no specific single criterion that exists for delineating the presence of significant interference. What do we mean in the product insert, what is a significant interference? There's no consistent definition of that.

Guidelines have been established for evaluating of interference effects, and we've heard some of those. There is the CLSI EP7. IFCC has a guideline. Several individual groups, individuals

have published in the literature different processes for approaching how we evaluate interferences and limitations for glucose monitors.

There is no current consensus exists among manufacturers about the most appropriate way to publish guidelines on how a interfering substance is affecting a particular method. We've heard from Dr. Ginsberg the idea of standardizing. And I'm all for standardization, and so we do need to look at how we can standardize the information that's provided to the end user.

And some other issues and questions: Should there be a standardized procedure for evaluating interferences? Yes, I think there should be. How aware of these factors and interferences are the end users, and has such an awareness survey been done? Do we really know how aware the end users are, whether they be in the ICU or whether they be the home users? Are package inserts enough? Are they adequate in the information that is provided to the end user? Certainly is more public education needed? And I think we'll all agree that yes, there is more education

needed.

Sources of information for the end user:

again, package inserts and instructions, are they adequate? Probably not. Where can the public get information about their blood glucose meter? Well, some of them go to the pharmacist to ask the pharmacist at their local pharmacy. These individuals are not trained necessarily, and aren't staying on top of the information, but this may be a source where many of the end users go to ask for information. If it's a hospital staff, obviously, they have their clinical chemist in their central laboratory that they can get information.

And the place where most people nowadays go get information, the Internet. And what websites are available? One website, and example, is the ACC's Lab Tests On Line, which provides end users with information about the glucose test and what the test means and what the limitations of those tests are. So there are websites that are available, but we need to look and see if these websites can be improved.

So this gives you a little bit of a

background, what I'd like for you to consider as we hear and talk more now in this session about interferences and limitations for blood glucose determinations.

So with that introduction, I'll introduce our first speaker in this session, and that's Dr. Courtney Harper. And as you know, she is the current Chief of the Division of Chemistry and Toxicology and Devices. And Courtney's going to talk about the FDA perspective, public health notification, potentially fatal errors with the GDH PQQ glucose monitoring technology.

DR. HARPER: Thank you, Gary.

So, today I'm going to talk to you a little bit about interfering substances. And I'm also going to talk to you about how FDA sort of tries to detect problems with devices. So this morning we heard about some inherent problems with devices. I'm going to talk a little bit from a different perspective about those problems, first of all how FDA collects that type of information, and second of all how FDA and others react to it. And what I hope to gain out of

this is by using one particular example, later on this afternoon I'd really like to hear input from the panel and also any of you in the audience about how can FDA and others better send messages and collect information about the problems with products.

So first, I'd just like to reemphasize. We've been talking today about a lot of issues with glucose meters. We keep talking about how they need to be more accurate, they don't do well enough in the ICU, or they do do well enough at home, or whatever we're talking about. But these particular devices, these types of devices, have been very beneficial. And we've heard a little bit about this, too. So let's keep in mind the context here, which is that without self-monitoring of blood glucose, diabetic patients would be worse off than they are today. So even without our current performance, I think we need to recognize that these are good products to have out there, and they have revolutionized the treatment and monitoring of patients with diabetes.

Most of you probably know the majority of glucose meters that are used at home use two different

types of methodologies. One methodology is the glucose oxidase method. You've heard some about that today. This is the one that's more oxygen sensitive.

The other type of methodology is a glucose dehydrogenase enzyme, and this particular enzyme can use three different types of mediators: PQQ, FAD, and NAD. And those actually have some separate types of interferences.

These types of technologies are used, as you all know and we've discussed, at different places. They're used at home by lay users. They're used in healthcare settings, and those settings vary quite a bit, from hospitals and ICU units all the way to emergency response units and also long-term care facilities and nursing homes. And perhaps each of these facilities and each of these use settings has a lot of different issues, as we also discussed today.

So hopefully the majority of the time, things go well, and these devices are used well and the majority of the time people get accurate results, or accurate enough results to use for dosing insulin or any other treatment choices they may make. But

sometimes things go wrong, and sometimes we have adverse events as a result of incorrect test results or the use of medical products. And FDA actually collects adverse event data from our medical device reporting system.

The medical device reporting system has a method for users, laboratories, and clinicians to actually report to FDA any problems they may have been with any medical devices. So if you take glucose as an example, if a lay user realizes somehow that they have an inaccurate measurement on their glucose meter, they can actually report that to the manufacturer, who then can report it to FDA. Or you can actually report to FDA directly. At the end of the talk, I'll actually have a link so anybody who wants to can actually report adverse events to FDA, and we encourage that they do so.

And when we get that data, FDA tries analyze it for trends. We look to see whether or not there's a trend of a problem, or we look to see whether or not there's a problem that's sticking out as something that needs to be addressed, or perhaps something

that's showing that things are going well or things like that. We can do that with greater or lesser success depending somewhat on product type and also the type of reports that we get.

So the medical device reporting is actually sometimes difficult to analyze, and difficult to find those trends. And this is particularly true for over-the-counter devices. So you have to keep in mind that blood glucose meters are used -- there are billions of test strips sold per year. So the denominator is huge, and the number of tests that are performed is huge. And we know that events are underreported, but they're also underreported in specific populations. They're underreported by lay users. They're probably more likely to be reported when there's a problem noticed at the hospital. And we know that overall they're underreported, and that even when there is a device result, either it's unrecognized that the incorrect result happened, or it's not reported to FDA. So the actual reporting of blood glucose meter adverse event data is actually a little bit misleading for us.

Another issue is that the data fields that we receive is often descriptive. So what happens is, depending on the reporter, sometimes there's more or less information about the event that's included. And this actually really limits the ability for us to do electronic analysis of large datasets. So we know there are some cases where you can search for a particular word, but if it was entered with an incorrect spelling or if it was entered with a different word or a synonym for something that's not often used, sometimes it's very difficult to group similar or like issues together.

So these are some limitations of the database. Now, it stands out for glucose meters, because if you have a product where perhaps there's only three or four of these devices implanted into patients in the U.S., you're going to be able to look at very closely every single adverse event report that comes to you, so you'll be able to read them individually. You can do a very good analysis. We get more than 12,000 reports on glucose meters per year, and so it's actually not feasible for FDA

analysts to read every single report and somehow group them or analyze them and call and contact every single reporter. So that's where some of the limitations of the database come through.

Now, the fact is we get these 12,000 reports a year, but also keep in mind that that's almost a little bit good, because we do have billions of tests performed per year.

Now, here's an idea of what I mean by how it's difficult sometimes to group some of the device reports. So we did a little bit of an analysis of the database, where we actually pulled out all of the glucose serious injury reports from 2004 to 2008. And I will point out that probably we would get different numbers if we used different search terms, so this is another limitation of our database. So for this particular search, we pulled out 12,672 serious injuries.

Now, each of these injuries is described in different ways. So here I've shown you the top 11 injury codes that came out of this search, and you can tell that the numbers don't add up to 100 percent,

because sometimes these things are overlapping. Also, the terminology that's inherent in the FDA database is confusing, because they have blood glucose low and hypoglycemia as separate events, and sometimes things are reported in both, and sometimes things are reported in one. So it very much makes the job of FDA in terms of looking for these trends difficult. So we're actually actively working on trying to improve some of the data analysis and reporting techniques for glucose meters over time to see if we can get some better data to try and make sure that if there are trends or events, that we can pick up on them.

What's a little easier to look at is the death data, because fortunately, there is a manageable number of deaths reported to FDA. These are also underreported, but like I said, given the volume of testing, 100 deaths since 1992 is actually not good in that you don't want any deaths, but it's better than having 12,000. So we actually did look at all 100 deaths between 1992 and 2009 in a single analysis, and we tried to classify them. And as you can see, because of the database, the majority -- or not the

majority, but the largest category that we had to put things in was "unknown cause." And that is actually a feature of our reporting database, where you might read the description of the event and you can actually not attribute any cause to the death that occurred. So you don't know whether it was associated with the meter or not, or you don't know what happened with the meter to make that happen.

Meter malfunctions would be 11. An example of this would be meters that inadvertently switched from milligrams per deciliter to millimoles per liter. If somebody might take an action on a number for millimoles per liter, thinking it's milligrams per deciliter, they may actually make an incorrect treatment choice.

False high results and diabetic ketoacidosis were also high numbers. And there were 13 deaths from the issue of maltose interference for glucose dehydrogenase PQQ meters.

And so I'm going to take this example to walk you through a case where we have identified an issue over time and the way FDA has gone about trying

to address this issue, because later on, if you have an input on how FDA can better do this to actually help put out the message, we'd really like to hear it. So I'd like to use this as an example.

So the particular issue with glucose dehydrogenase PQQ is the following. So glucose PQQ -- in the interests of time, I'm not going to try and pronounce PQQ for you, but there it is on the screen. This is a type of technology that's been around for a very long time. It's been marketed for over 20 years. This technology is non- selective for glucose, and that means that it also detects other sugars, such as maltose, xylose, or galactose. And these meters were actually already around when certain drug products were approved by FDA that contain these sugars as components. And when that happened, sometimes when patients were treated with these drugs, the meters actually detected those sugars as glucose, to some extent, and so although the meter is actually giving a correct reading of the sugar that it's seeing in the blood, all of it isn't glucose. So taking an action on that number can be inappropriate clinically. And

there have been deaths and serious injuries associated with this, and also severe hypoglycemic events.

Now, both the devices and the drug are labeled against this use, so there's clear labeling on the devices and clear labeling on these drugs, as well, to say that these two products shouldn't be used together because of this interaction. But as I'll show you in a minute, some of that hasn't been effective.

So the types of drugs -- I'll list them more completely later. But a lot of them are IB drugs, because that's where you get enough exposure to these sugars to actually cause the issue with the reading. So Extraneal would be a peritoneal dialysis solution that's been involved in some of these issues, and also some IB immunoglobulin solutions that use maltose as a stabilizer.

So FDA actually recognized this problem several years ago and has taken several actions over the last several years. So in 2005, we released a MedWatch Safety Alert on this issue. In 2006 we did another Patient Safety News to try and reach out to

healthcare providers to describe this problem. In 2008 we published two articles, one out of the Center for Drugs and another out of the Center of Biologics on this particular issue, describing deaths that may occur, and also another Patient Safety News on the topic. And last spring, in 2009, the Center for Drugs actually updated the package insert for the Extraneal drug product to include a back- box warning against using GDH-PQQ test strips on these patients.

Regardless, from 1997 to 2009, FDA has received 13 total deaths reported to FDA associated with this glucose test strip. We've actually seen a few other cases in the literature that weren't reported to FDA, as well. These deaths all occurred in healthcare facilities, and six of the 13 deaths have occurred since 2008. So this told us that even though we've been doing sort of outreach efforts for the last several years, these deaths continue to happen at some rate, so it really gives us a signal that some of our outreach efforts are not being effective enough.

Ten of these 13 patients reported were on

Extraneal, the drug that received the black box warning, and three of the 13 patients were receiving other maltose- containing substances, like these IV immunoglobulins.

So as you can see, we have a problem. These previous actions that we've taken are not having an effect on deaths, and we are actually continuing to see deaths. It's almost as if when an alert goes out, there's some period of awareness and then there's a little bit of a decline in awareness, perhaps because of turnover of workers in healthcare facilities.

And the other part of this problem is that this is an issue that doesn't actually affect the majority of patients using these devices. So this is a very -- this is a pretty small minority of patients that are on these devices. But when this happens, it's devastating for that patient, so it is a very serious problem when it happens.

So FDA, we thought that additional action was warranted, but what we were worried about, because this affects a small minority of patients and because testing for blood glucose in diabetic patients is so

important, we really struggled with how to balance the message, to make sure that people continued to test and did not misunderstand the message and believed that there was a problem that applied to them, when really we wanted to reach a very small population of patients and the healthcare providers who treated them.

So how do we balance a safety warning and, you know, getting the message out that patients, you know, really need to keep testing? So the communication of complex issues is definitely something that we really work on and try to do well, and I think sometimes we succeed and sometimes we don't.

So this time we actually decided on a stronger message, hoping that it would get a little bit more attention in the healthcare facilities, and also some ongoing other actions. So FDA published in August of 2009 a public health notification on this issue. And the point was to try and raise the level of the recommendation to attempt to increase awareness about the problem. And this public health

notification was aimed at healthcare providers and healthcare facilities. And it described the nature of the problem in detail in the public health notification, and the link was on the previous slide. And it described that this issue, when it's unrecognized and when these test strips are used on patients that have these drugs in their system, may lead to inappropriate insulin dosing that could lead to serious injuries, such as severe hypoglycemia or perhaps leading to coma and also death. And also, it can lead to unrecognized hypoglycemia. And it can occur anywhere, so although these -- all the deaths have occurred in healthcare facilities, it is possible for some of these drugs, which are outpatient drugs, that it could happen at home.

We also tried to describe that this is not a problem with other types of technologies. It does not affect glucose oxidase test strips. It doesn't affect the other two glucose dehydrogenase test strips, which aren't maltose-sensitive. So the GDH-NAD and GDH-FAD have a much lower sensitivity to maltose, galactose, and xylose. And it also doesn't affect laboratory-

based methods, so usually those methods use hexokinase, which is a little bit more sensitive.

So the recommendations that we gave to healthcare facilities were these. We actually recommended that healthcare facilities avoid, in general, the use of glucose dehydrogenase PQQ test strips in their facilities. For those facilities who continue to use, or in the meantime before they switched over, GDH-PQQ test strips, we recommended that they never use them on patients receiving these particular drugs or products and/or on patients from whom they couldn't get that information, so patients who were perhaps unresponsive when they entered the hospital. And we recommended that only laboratory-based assays be used on these patients.

We also recommended that not only hospitals determine whether patients are receiving these upon admission, but also periodically through their stay; that hospitals increase efforts to educate staff about this issue; that they consider using drug action alerts in their hospital information systems, and that they periodically verify glucose meter results on any

meter with laboratory-based methods. And the public health notification also lists the particular drugs.

As I mentioned already, Extraneal and some IV immunoglobulins -- Orencia, Adept adhesion reduction solution, Bexxar(ph). And also we included a bullet saying any product containing or metabolizing into maltose, galactose or xylose.

We have had some questions on this, because it is a little bit vague. And the reason it's vague is that there are some products out there where the content of maltose or these sugars isn't known, or there could be compounding within hospitals using some of these products.

So what happened in some of these cases is that you have an issue where a test or result on one of these strips and the patient who receives some of these drugs might be three to 15 times higher than the lab result. For example, one patient where the blood glucose result on the meter was 200, the lab result at the same time was 19. And when those patients are treated with insulin, it drives them even lower, and that's where the deaths have occurred.

So at the same time we released the public health notification, we also released an advice for patients. This is a publication that accompanies the public health notification and is intended for lay users. And it held similar, but separate recommendations for these lay users, because we were trying to balance the message to make sure that people could find out whether they were affected or not and what they should do in the meantime.

So we recommended that diabetic patients who any of those blood products should never use test systems that use GDH-PQQ test strips, and that they contact their healthcare provider if they aren't sure if they have those particular types of technologies or are on those drugs, or if their results don't reflect the way they feel.

We went on to make general recommendations for all diabetic patients that they continue testing; that they not change test strips that were intended to use for their meter; that they try and find out or understand the type of technology they're using, the drugs they're on; and to make sure that they know that

any meters and strips that don't use this technology are okay for them.

And in the meantime, we're still working with manufacturers to address some of these issues. And we're also monitoring the adverse event reports to see if this problem keeps occurring. In the meantime, like I said, we encourage all facilities and users to report problems of this type or any type to us so that we can continue to do this type of trend analysis.

Because here's an example that I've given, where you have an interference that's known. So this is something that a lot of people in this room probably already know about. This interference is predictable. If you're on one of these drugs, you know that this is going to happen. And it's preventable, because if you know it's going to happen, then you can certainly avoid it.

But the problem is that the awareness has been too low, and it continues to be too low. So we're trying to figure out what is the right balance of how much is enough to do in this case.

So our challenges that I'll leave you with

are that past communications have not had a lasting effect, so we would like to hear from you all about potential other ways that we can do outreach and education on issues like this or whether other actions are necessary to mitigate the risk of this type of problem. And what more can we, the healthcare industry, and the medical device industry do to prevent some unnecessary deaths due to known interferences like this and like other interferences?

So I thank you for your attention, and I look forward to the discussion. Thanks.

(Applause)

DR. CLARKE: We have time for one question.

MR. KIECHLE: I'll be quick. It's Fritz Kiechle, clinical pathologist from Memorial Healthcare System in Hollywood, Florida. Excellent presentation. This sort of represents a never event, and it's clearly a patient's safety issue. But it's one that a hospital is perfectly capable of tackling if done carefully and with diligence and surveillance.

And the way we've handled it, and I've dealt with this problem in two different hospitals -- we

handled it the same way both places. Those items that contain maltose, for the most part, the interfering substance, are in the pharmacy. The pharmacy has a committee, has a formulary, which controls the substances they carry in the pharmacy. And they have a Pharmacy and Therapeutics Committee that makes those decisions. The first thing we did was go to the P&T Committee, as it's known, and asked them to remove these maltose-containing substances from their formulary and find substitutes. If that couldn't be done or there was an exception where there might be a physician who insists on using one of these compounds, we would then send the compound with a note from the pharmacy reiterating that we will not use this particular glucose meter for these patients while they're taking this substance. And we also get a phone call to the point of care coordinator the minute the drug is shipped to the floor. That person goes up and re-instills with the actual nursing staff that are taking care of the patient that they should not be using the glucose meter.

And that really has worked quite well, and

as far as I know, we've had absolutely no problems.

DR. HARPER: I'm really glad to hear that. I mean, that sounds like your hospital has taken this problem seriously and tried to address it. And that sounds like it will be effective when you are actually the ones administering the drug. Some of the reports have actually been cases where a patient had had the drug outside of the healthcare facility in which they are being treated, and they actually then arrived at the healthcare facility and in some cases told the nurse or nurse practitioner that they shouldn't use that meter. And unfortunately, there wasn't enough education of the person running the test to know that maybe they should believe the patient. So it actually can be quite tragic.

But I'm glad to hear that there's a lot of risk mitigation procedures being put in place.

MS. SKEENS: Can I just clarify? This is Lisa Skeens from Baxter Healthcare. And I just want to point out that for Extraneal, where this a very serious issue, as you've pointed out, that maltose actually stays in your body for two weeks after a

patient is administered. It's not prescribed by the pharmacy, so that actually would not mitigate this risk for Extraneal patients. So it's, as you mentioned, very important to have other types of education involved, because it's not your typical situation, and Extraneal is a chronic use product. So again, a patient may be administered not receiving the product in the hospital, but have this serious issue.

DR. HARPER: Right. Thank you.

DR. CLARKE: Okay. Thank you again.

It's time for our break. We'll take a 15-minute break. We'll reconvene at 3:20 or when you hear the bell ringers in the crowd.

(BREAK)

DR. MYERS: If I could ask everyone to take your seats, we'll continue with our second session.

Okay. I want to welcome you back to the continuation of our second session, Blood Glucose Meter Performance, Interferences and Limitations. Our second speaker today is Ken Ervin. And Ken obtained his bachelor's and master's degree with a focus on analytical chemistry from the University of California

at Riverside. He's held positions with R&D with Syva Corporation, Abbott Diagnostics, and SmithKline Instruments. In 1983 he joined the startup company of LifeScan as an R&D director, and there he participated in the development of Glucoscan, One Touch, and SureStep blood glucose monitor systems. And he took early retirement in 2004 and has been consulting primarily with startup companies in the blood glucose field.

His talk today is going to be Analytical Interferences and Physiological Limitations of Blood Glucose Meters. Ken?

MR. ERVIN: Thank you, Dr. Myers. As you may have gathered from that introduction, I've spent a good part of my career within the glucose monitoring area, and as such, I have observed evolution of the product over this time frame.

Initially, the biggest place to hit accuracy was with user error. And a lot of the early products attempted to do that. We ended up with what were called second generation products, where we did away with that wiping and blotting business, and that was

the major source of user error.

But we've also seen evolution in the technology itself to provide a more convenient, faster, and cheaper product. And we are also seeing evolution of the technology to produce more accurate results and to deal with some of the lingering problems. One of these is the area of analytical interferences and physiologic limitations.

There's a plethora of information out there on this topic. And if you're really interested in it, I think two reviews in particular by Wahl and Duncan are particularly useful, in that they're pretty comprehensive and pretty informative. There's also a number of original articles. I've just got partial lists here in both categories. There's no shortage of information.

We've already heard about package inserts several times today. And generally, these things are captured in the section that's typically called procedural limitations. And they're bucketed into several different categories: things that are considered to be sample- related, that is, the sample

hematocrit; po2; whether or not the patient's in DKA or HHNK. There's a category for endogenous compounds, sort of naturally occurring substances which, should they become elevated, could possibly create a problem; exogenous compounds -- we just heard a bit about maltose, but there are others; and of course the environmental considerations.

Now, one of the things we were talking about a little while ago was accuracy. One of the main differences between blood glucose monitors and laboratory instruments is -- I know this coming from a laboratory instrument background -- they do everything they can to control things like temperature and the environment that the test is being done in. Blood glucose monitors probably could do that, but as we've heard earlier, the cost of the devices would probably go up exponentially. It takes a lot of power and a lot of work to do that.

So what they've done is to sort of confine the use range in terms of temperature or in terms of humidity. Altitude is essentially a po2 problem.

One theme that I'd like to carry through

today is that each manufacturer, in producing their product, has some design goals that they're going to go after, and that those design goals to some extent, and also the effort to achieve a proprietary position, they're going to pick some set of technologies to put into their system and give them a competitive device. The choices of those technologies in essence leads to some of these interferences and these limitations.

However -- I'll go back to the point I wanted to make earlier -- technology is evolving, and you are seeing products in the marketplace now that are dealing with some of these, and dealing with them quite effectively.

This is just a partial list of some design goals that a blood glucose manufacturer might consider. The obvious ones, accuracy and precision, specificity, et cetera. I've asterisked some very important ones from their perspective, though. The product has to be stable. That is, the test strips have to be stable at room temperature. We've already heard some discussion this afternoon about the impact of using fresh and late-dated strips.

The test has to be rapid. Now, we're talking in the order of five seconds now for many products that are available to the consumer. That basically -- and the environment where we're dealing with it in the home, in particular, means that you're going to use whole blood. You don't have time to deal with plasma or serum. So the use of a whole blood sample is a key piece of the design goal.

And then the second -- or the third item here I've asterisked is, it has to be very easy to use. If these products are going to be used even by nurses in the hospital, they've got to be easy, because they don't have time to pay attention to a lot of technical detail.

Another item I'm going to mention here, because it relates to accuracy, is you've got to develop some sort of calibration strategy. And historically, the calibration codes have actually been windows. It could be two, three, four percent that a particular cal code would represent. And so one of the areas that manufacturers obviously are trying to work is how they can actually pull these calibration

codes into tighter and tighter ranges.

So in order to meet these design goals and specifications, they choose particular technologies, both for the device and for its method of production. The other thing that's worth remembering is that blood glucose measurement is based on combining technologies. For example, you have to have a method of getting your sample into the device. And currently, most methods are using some sort of capillary action. There are some that are using, for example, blood drops on a membrane, but for the most part, everybody's headed toward some sort of capillary approach.

You have to have a method to identify glucose, and you want to do this without question. This is the reason for the GDH-PQQ problem in the maltose. It's referred to as specificity, and in fact, there are three enzymes that have been used in blood glucose monitoring devices. All of them are fairly common in the laboratory, and they've all been used in blood glucose monitoring devices, as well.

And finally, you need to have a method to

quantify the glucose. Essentially right not, it's either a photometric, colorimetric assay or an electrochemical assay. You also have a method of calibration, which I mentioned a moment ago, and this not only relates to sort of the samples that you have to use, the reference system that you're going to use, the width of your cal code space -- a number of important variables. And in recent years you've seen more effort in the area of assessing whether or not a test was done correctly. Is there something that could have gone wrong? Was the temperature incorrect? Instrument manufacturers have been trying to evolve their technology to identify these conditions, and even more recently now, if they've identified a significant contributor, can they correct the results for that?

So interferences result in two areas. One, from the analyte specificity or some effect on the reaction that follows, either a sample influence or an environmental influence on the measurement reaction. I've said the enzymes that are being used are glucose oxidase, glucose dehydrogenase, and hexokinase.

Glucose oxidase in the original products was exclusively used. All of the original blood glucose monitoring products used that, for two reasons. One, it's very stable at room temperature, and secondly, it was very specific. It could easily be formulated into a test strip. And then thirdly, it was easily coupled to a colorimetric indicator system, of which there were many, and you had your blood glucose monitoring device.

GDH has historically been most used with the HemoCue system -- I mean, that's the one that's been out there the longest using GDH. That enzyme is particularly insensitive to interferences. However, it does have a tradeoff, in that it's not particularly stable, and therefore their cuvettes have to be refrigerated.

There was one product from Bayer, called the Encore, that actually used a hexokinase G6 PDH system, affording it great specificity. And I'll talk a little bit more about it later on when we're talking about po2.

We've heard a lot already about endogenous

substances and exogenous substances, but what I really wanted to talk a little bit about today is not what they were and how much effect they have, but rather, how does that happen? Let me take ascorbate as an example.

In a colorimetric system using glucose oxidase, where we produce hydrogen peroxide as an intermediate product, and then a final colored product, ascorbate has the property that it can either reduce the peroxide or the colored product and actually produce a low result. In some electrochemical systems, it can be read independently of the glucose reaction and produce a positive bias to the result. Fortunately, the electrochemical systems are now beginning to incorporate, as Dr. Ginsberg has mentioned, schemes to identify the occurrence of this background current and to eliminate it from the glucose calculation.

When we talk about patients with DKA or HHNK, I don't think I've seen very much about a pH problem with glucose test strips. That's not to say I don't think they've happened, but generally the test

strips are highly buffered so that pH shifts in DKA are not going to be much of an issue. On the other hand, viscosity of the sample -- remembering that these samples tend to be dehydrated, and as such, there's very little water left in them -- they tend to be hyperosmolar. And these can have flow effects in a system, particularly in capillary systems.

Back in my days at LifeScan, we had an HHFK sample that I put on top of a test strip. It took over 30 seconds to penetrate into that test strip when a really high hematocrit sample might take only three to four seconds. There was so little water in that sample that it wouldn't penetrate the strip. When it did, it produced a very low result. So these kinds of samples can have what I call flow effects on your glucose monitor. They don't happen very often, but that's why they're in the labeling.

Environmental influences -- I mentioned earlier that laboratory instruments are trying to control everything. Glucose monitors just try to control by defining a range in which they work. The altitude thing is basically a po2 problem.

When we go to looking at physiological limitations, we basically have three choices of samples, or after talking with Jeff DuBois earlier, I probably should have said "specimen" here -- capillary, venous, or arterial. This presents us with three or four different difficulties. One is that the actual concentrations of glucose can be different in the three, and that's well understood. We've also heard today reiterated several times that hypotensive patients, other conditions creating profusion problems, and I ran across a patient once where Reynaud's Syndrome actually created a big difference between the vein of puncture and the capillary result. We've heard already today about the alternate site time lag and the po2 differences.

Okay. What I wanted to do today, and since I'm beginning to run out of time I'm going to have to hustle along here, I was going to cover two examples, hematocrit and the po2 problem. These reactions on the screen, the upper one is that typical glucose oxidase reaction where oxygen is used. And this is where the po2 effect comes from. In systems where

oxygen cannot get to the sample, or where the sample itself varies a great deal in its concentration, you can create an effect that is the result of competition. So for example, in an electrochemical system, where they use mediators, because they're more efficient in turning over the current, the - - this reaction will be in competition with the glucose plus oxygen reaction. And we get a competition now such that in the case of venous blood, you're going to get higher results, because there's less oxygen present in the sample. More of the glucose can go the mediator pathway, and so you tend to get higher results. Arterial tends to read lower, because you have a lot more oxygen present. By the way, this is all relative to being calibrated with capillary blood, so you have more oxygen, or in the case of highly oxygenated venous samples. You can take a venous sample and mix it with air for a while, you'll get the same effect. So we have a competition going on there that drives this po2 effect.

The second-generation products had a similar difficulty if their reaction site was somehow occluded

from exposure to the atmosphere. Products like One Touch and SureStep, where the membrane surface is open to the atmosphere, they didn't have much of an issue. But other products, where they use a window, you no longer could get oxygen in there, and so you ended up with this po2 competition problem.

I mentioned the Bayer product earlier, hexaconagesic (ph) PDH. They went with this system ostensibly to get away from the po2 effect, but they also had some stability issues, as well.

The GDH PQQ system was introduced to alleviate this po2 effect. Doesn't involve oxygen, so there's no competition. And this particular enzyme happened to be relatively stable at room temperature, so it had that attribute. However, that particular enzyme has demonstrated less specificity for glucose, as we've heard, and recognizing other glucose-containing sugars. And hence the problem.

In terms of the evolution, people have been working on trying to develop GDH enzymes that, again, use the NAD or FAD co-factor. And these are, through their genetic engineering, becoming more specific and

more stable, and you're seeing some of these in products now.

The other condition I wanted to talk about was -- that we've heard a lot about today is hematocrit. One of the problems here is that because we were dealing with so much confusion between understanding the difference between whole blood and a plasma result, a lot of the manufacturers moved over to reporting plasma-equivalent results. These systems have to be calibrated at normal hematocrit, but we know that the hematocrit ranges can vary substantially, five-fold, or nearly five-fold. So applying that calibration to samples that are significantly different than normal is a source of error.

In this illustration, what I'd like to show, the blue line is what you would expect just on the basis of physiology. In other words, if you take a whole blood sample and plasma from that same sample and look at the glucose content, as you go down towards a sample with a hematocrit of 20, you're going to have basically about a six percent bias. If you go

to higher hematocrits, you've going to have a six percent negative bias. That's just the physiology of that sample.

And as Dr. Ginsberg mentioned earlier, hematocrit is a more complex issue, and he already alluded to the business of red blood cells and how they may impact the measurement itself.

The yellow line illustrates a product which has minimal what I'd call method effect related to the technology, the green line representing a product which has increased or greater effect as a result of its technology. And some products are really sensitive, to the extent that as you approach higher hematocrits, they just nosedive, and that's illustrated here by the red line.

And as he was suggesting, hematocrit does influence the access of plasma or the diffusion of glucose through that sample in the test device, suppressing results. So you would expect, then, that when you're really trying to turn over reactions quickly, like at higher glucose, the effect is going to be greater. And in fact, that's observed. If you

look at some of the labeling and products a while ago, the -- they actually had two different hematocrit ranges, one for lower glucoses and one for higher glucoses, the higher glucoses being somewhat compressed.

The good news is, and several companies are doing this now, it can be measured, and it can be corrected. One question I have there is, is there a price? Now that you're making two measurements, is there a price you pay in terms of imprecision?

So in conclusion, limitations and interferences are related to the particular technologies. The goals of a BGM system to me make it unlikely they'll ever completely match a lab-based system, but they'll get close. They can get close. And the evolution of these devices is a demonstration of trying to achieve this balance between a high degree of performance with a rapid, versatile, easy-to-use system.

And of course, the last conclusion is when we're talking about hematocrit, we've got an automatic plus or minus six percent.

Thank you.

(Applause)

DR. MYERS: We have time for some questions.

Ken, I have a question for you. How many different bases of calibration for the instruments are you aware of, as far as methodology calibration?

MR. ERVIN: Oh, one of the reference bases early on was YSI, because it dealt very easily with whole blood samples. So a lot of the instrument manufacturers used YSI initially.

Also initially, all of the testing was intended to be done on capillary blood, because these were intended for users at home. As we moved to trying to improve accuracy and the versatility of the product, you started seeing other reference systems utilized. For example, one company I know used a deproteinized hexokinase as their reference system.

If you're asking the question about the calibration codes themselves, I don't have information. A lot of it's proprietary. But there are very different schemes in the way people approach calibration of their devices. Some will use equations

built into the system to actually describe the behavior versus whether it's a current or a photometric measurement. Others will make sure that their product behaves in a linear fashion, and then they just assign a cal code that normalizes their result to their reference value.

I'm not sure if that completely answers your question, but ....

DR. MYERS: Well, I guess the question is, some are, as you say, calibrated to the YSI, and some are actually calibrated directly to mass spec. So is there an issue among meter variability because of how the variation in what the calibration point is?

MR. ERVIN: I hope I didn't say GC mass spec. I don't believe any of the manufacturers are using that as their routine method for calibration. When you're in a manufacturing situation, and I'm sure they're going to address this themselves, but when you're in a manufacturing situation, you've got to have something that is feasible for routine day-in and day-out operation. And I'm not privy to what all of the different manufacturers are using as their

reference systems. For example, Nova probably is using something very different.

But initially, and it's evolved from that, but initially everybody pretty much was using YSI, because of its convenience in dealing with whole blood samples.

DR. MYERS: Thank you.

Okay. Our next speakers will be two individuals representing the industry's perspective. And I'll go ahead and introduce them both at the same time.

First of all is Dr. Alan Cariski. And Dr. Cariski received his bachelor's degree from John Hopkins University, his medical degree from the University of Rochester School of Medicine and Dentistry, and his law degree from the University of Maryland School of Law. Dr. Cariski has over 30 years of experience in the healthcare industry, having practiced medicine as a board-certified internist and endocrinologist before transitioning to industry in 1991. Since 2001, he has been Vice President of Worldwide Medical Affairs and Medical Safety Officer

at LifeScan.

And our other individual is Mike Fliss, and he has a BS degree in mechanical engineering from Marquette University. His industry leadership experience includes serving as Co-chair of AdvaMed's Blood Glucose Working Group since 2007, and he has held a variety of key regulatory roles during his 23 years at Roche Diagnostics working with product developers to optimize product design characteristics, product risk assessments, clinical trial designs, and instructions for use in promotional materials to facilitate prompt registration processes.

Gentlemen?

DR. CARISKI: Thank you very much, Dr. Myers.

Why don't you manufacturers make glucose meters more accurate? That's a question a well-known endocrinologist asked me at the Diabetes Post-Graduate course in San Francisco that was held recently. And as luck would have it, I didn't have this slide set with me. But I have it today, and I'd like to spend the next few minutes answering that question for you

as best I can.

So the glucose manufacturers as a whole are committed to designing and manufacturing glucose meters to meet the needs of individuals with diabetes.

And industry shares a goal of advancing meter technologies and improving accuracy through innovations in meter systems that accomplish these three things:

Reduce use error. And I think it's important that most people today have used the term "use error" rather than "user error," because the term "use error" recognizes that it's not only the user, per se, but it may be the way in which the product was designed, so that any user might encounter problems. So I think use error is an important term.

Also, the manufacturers try to reduce the impact of interference and improve the overall quality of testing for patients.

And finally, both the glucose standard for consumer meters, that 15197, and industry recognize the role of design to improve not only analytical performance, but also the usability of the instruments

to increase patient compliance with glucose monitoring regimens.

It's instructive to compare the meter to the other technologies that are available, and to some extent, Ken and some of the other speakers have done that. The so- called definitive method uses isotape-labeled glucose dilution mass spectrometry with gas tromatography. And while this is something you might think that no home should do without, I think you have to concede that it's not something you're going to take on the road with you. The method requires meticulous and time-consuming serum sample preparation.

The same thing is true for the reference method. It's time-consuming, requires serum or plasma. And finally, the typical diagnostic laboratory method also uses prepared serum or plasma, and the environment is very precisely controlled, both the environment in which the device sits and also the environment within the device itself in terms of pH reaction initiation, et cetera.

The meter, on the other hand, uses capillary

whole blood, a point that Ken just emphasized. And it emphasizes ease of use. It has to be portable; it uses small volumes, people want fast testing. And it has to be used in a variety of environments. The lab, it has, of course, a controlled environment. There are restrictions on where meters can be used, but they're fairly broad in terms of temperature, altitude -- most will work up to 10,000 feet -- and most have a fairly broad range of relative humidity.

And of course, these tests are performed by lay users, not by trained laboratory technicians. And when you think about it, we're really asking lay people to perform a laboratory test; okay? And we're trying to make it as simple as possible and still be accurate.

So the next few slides will compare and contrast the lab instrument to the glucose meter. So, as I mentioned, the lab has definitive reference methods. The glucose meter can at best say that it's ultimately traceable to the definitive or reference method.

There are standard reference materials to

test the lab instrument. These don't exist for glucose meters.

The hematocrit effect we discussed. In glucose meters, the hematocrit effect is mitigated by measurement, but not necessarily entirely, and also by the use of algorithms.

Finally, there are some fairly accurate devices that use whole blood, but they eliminate the hematocrit effect through membranes, and the membranes also protect the enzymes and sensor from interfering compounds. SMBG technology, at least what's available today, doesn't have these advantages. And the presence of multiple interference provide a considerable technical challenge.

Again, laboratory instruments tend to cost more than $10,000. A typical glucose meter is less than $100. Lab instruments require maintenance; glucose meters don't. Plasma versus blood -- giving of plasma equivalent. Trained technician versus a lay person -- the lab instruments are typically calibrated many times daily. Glucose meters cannot be calibrated by the user. The controlled versus the uncontrolled

environment I've alluded to. And in the lab, controls are run frequently. Control solution is available for users to test the integrity of their system. Many people don't use it, or use it sparingly. And if they detect a problem, unlike the lab, there's nothing they can do, at least with the product they have.

A laboratory instrument is large and stationary, susceptible to shock. The meters, you know, are brought everywhere. They have shock testing. They're very tolerant to shock. Most lab equipment requires at least five milliliters of blood. I expressed it as microliters just to contrast with the glucose meter, which requires less than a microliter. And the meters requiring the smallest samples today require just .3 microliters.

A typical lab instrument takes at least 60 seconds to give a result. The glucose meter, less than 10 seconds, and the majority of meters today take less than five seconds, or no more than five seconds to give a result.

The inaccuracy of a lab is generally in the range of plus or minus four percent. Studies have

shown some are as high as plus or minus ten percent, but I think in general it's closer to the lower level. As Dr. Ginsberg mentioned, in general the meter can be much more than two times as accurate as the laboratory instrument on which its accuracy as being based.

So this is a slide which, by the way, took me hours to produce.

(Laughter) But it shows the -- the basic reaction, and I'll just go over it briefly, as Ken has spoken about it a little bit. So glucose is converted by an enzyme with more or less specificity. There's sometimes co- factors. And that enzyme is an oxidizer. It produces electrons. The electrons often are passed on to a mediator, and the mediator captures and transports the electrons to a test strip electrode or a test strip indicator. The meter reads a current, which is proportional to the glucose, or a colorless dye precursor is converted to a dye, which is read photometrically, and the changing color is proportional, again, to the glucose concentration.

So that's the basic principle. Of course, there are various things that can interfere. We've

already mentioned things like inadequate washing, that milking the site can dilute the glucose. Various enzymes have interference. Courtney spoke at some length about GDH PQQ. The mediators are sensitive to substances such ascorbic acid, or Vitamin C. The electrodes are sensitive to a variety of endogenous compounds, such as uric acid. And despite all this, the enzyme mediator actions have to be fast to yield a fast test time.

So there are a number of sources of inaccuracy. Some are analytical and some non-analytical. So interference, we've talked about endogenous interference. The one that presents the greatest issue is hematocrit. And then there are exogenous compounds, such as acetaminophen, the environment, temperature, relative humidity, altitude.

Misuse can be exposing test strips to high temperature and relative humidity. It's very hard for manufacturers to prevent a test strip that's been misused to be -- for a patient to use that test strip. All you can do is label.

And industry carefully considers all these

factors and the customer requirements, for example, hospital versus consumer, in providing devices that deliver maximum medical benefit.

Another important thing to remember is that the accuracy and precision testing of individual test strips is destructive. So any individual test strip has not been tested; okay? The lot has been tested. And the release criteria rest on sampling and statistical modeling. But to my knowledge, no manufacturer has a technique today that allows him or her or it to say that any given test strip meets specifications. It's a statistical statement.

Patents -- one of the reasons for the existence of patents is to provide blocking of other manufacturers from using your technologies. There are variabilities in raw materials that are not always understood, and small variations may have a substantial impact. It's incredibly difficult to tease out which ones to address to try to make the test strips work uniform and to extract whatever remaining variability in an individual test strip that is that remains.

And manufacturers are committed to meeting these challenges in ways that address patient needs as to cost, convenience, accuracy and precision, and data analysis.

So the challenge to industry is to manufacture high volumes of glucose meters and test strips. And in 2009 on a worldwide basis, individuals tested 17 to 18 billion times annually, which works out to 47 to 49 million tests a day. And the corresponding figures in the U.S. were 6.2 billion tests annually and 17 million tests daily. So that's a lot of product with which to maintain quality control. And any given manufacturer may make several billion test strips a year.

And I'll give the floor to Mike.

DR. MYERS: Thank you, Alan.

MR. FLISS: Whereas Alan is a medical doctor, I'm a little bit of a process geek, and it's my role to help my company register products. And I'm a big fan of national and international standards, because they create a roadmap. They're built upon consensus of industry, health authorities, and the

diabetes community. So they lay out -- if a manufacturer follows these requirements or the roadmap for developing a product, manufacturing a product, evaluating a product, and then presenting that data to the health authorities, they're like to receive a quick and favorable outcome to their registration, and that's helpful to the public that is served by the products.

Let me push through a few of these.

We've already heard today that there's an interest in addressing the performance characteristics of hospital systems differently than self-testing products. And industry supports that view, that not only from the perspective of accuracy, but also in the identification of potential interferences and limitations of procedure. We're working in conjunction with the ISO standard group, as well as the CLSI working group. And I wanted to clarify something that may have been said this morning. Neither one of these activities is close to bearing fruit at this point. The working groups are still gathering ideas and putting together their first drafts of the documents.

The CLSI committee, I believe, has two documents that they're working on. The ISO group is trying to revise the self-test blood glucose document that was created in 2003. So there is still quite a bit of an opportunity to have your voices heard. And the people who are involved in these groups, several of them are here with us today. In fact, Dr. David Sacks, who spoke this morning, is the chair of the CLSI committee that is examining what to put in this point of care document for glucose testing.

At a recent ISO meeting held in Washington, D.C., in January, we considered interferences and thought it might be best if we could create a master list of all the compounds that manufacturers should consider for self- testing and what concentrations those compounds should be evaluated at. And now the work -- it's a homework assignment a couple of the individuals accepted -- to put together a straw man proposal that could be shared with the other members of the ISO team. And that first document is going to be addressing self-testing.

From what we heard this morning, that

hospitals would have different requirements than self-testing, we'd like to see that perhaps the CLSI group, when they see what the ISO group creates for self-testing, can look at that and see, could a similar document be created for glucose tests that are performed in a hospital setting? And that document would, it would seem -- need to have a mechanism to be periodically updated. If we find that it's difficult to do that within standards, we might defer to a local health authority who has the ability to occasionally issue guidance documents, and in that guidance document perhaps as attachment, we could have this list of potential interfering compounds and the concentrations at which they should be tested.

What ideally -- if you can imagine, if there are over 20 companies that are competing to offer blood glucose monitoring devices, and as Dr. Cariski mentioned, there are almost 18 billion strips that are used over the world in a year. But only about a third of those are used in the United States. The other two are used globally. If 20 companies come up with 20 different ways of evaluating hematocrit, and all of

the interested health authorities around the world have a different way of looking for data from hematocrit, that is chaos. So if possible, within the next version of the international standard, we would like to see if we could address similar, as you see today, that there's a section that refers to accuracy studies; another section about user performance evaluation -- let's introduce another section that refers to how to evaluate potential interferences and how to present them to the health authorities so that they can make quick and informed decisions.

There was some mention by Dr. Ginsberg of reexamining how labeling should be laid out so that it's more helpful to our users, and I think that's an initiative that we all support. And I was reminded that back in 2001, FDA published a really nice guidance document for patient labeling. Now, it really goes into what you could say in a package insert. And what it -- if you're familiar with the reagents that come into the laboratory, they all follow a standardized format. That format is actually in the regulations. But for a self- testing product,

you're given some amount of latitude to try to take your content, your text, and write it in such a way that the self-tester would understand. So we're allowed to deviate from the -- the format of the centralized lab, and the companies are actually deviating from each other, as well.

So if you wanted to compare two products side by side by pulling out the package inserts from two different manufacturers, you might have a little bit of difficulty finding the similar claims. And this is something that perhaps we could work together through consensus as to the best way to lay out the package inserts so that people can get the information they need. And then Dr. Ginsberg's idea from this morning of let's perhaps get more information out at the point of purchase, perhaps on the carton label or on the vial label. Just as an idea, we know on the back of the carton label today, we have the range and composition table, which isn't very helpful to our lay public. So that's space that could be available to convey to the user what are some of the key limitations and procedure of that particular product,

and if there's a way to characterize accuracy with a number or small display.

So what we would like to conclude is that it is somewhat challenging to create many millions of these devices when we have to perform the tests with whole blood samples. We're aware that there's a need to evaluate several potential interferences, and we wish to work with the consensus developing organizations, both domestically and internationally, to create lists of potential interferences that should be examined and how they should be examined so that whether -- whatever product you're buying from whatever company, you can be confident that they all evaluated their product in a consistent manner.

As was mentioned earlier, it is challenging for these products, because they're used in a wide variety of settings. There's a potential the user might misuse the product beyond the manufacturer's instructions. And we have constraints for size, the meter sample, and the test time in order to make the product appealing so that people will reliably perform their tests and comply with their monitoring program.

So one of the key pieces we've been asked, how do we make our decisions as to the final design of our product? And we really try to take into consideration the requirements that have been expressed by our customers. And then we co-optimize across all these potential decisions we can make to design the product which is best-suited for that particular part of the marketplace.

Perhaps for the consumer channel, we might want to emphasize having a broad operating temperature claim so that we're confident that if there is a parent who wants to run a glucose test on a soccer field in Georgia in August, they're going to get a quantified test result, rather than an error message.

But we might take that same product and look at it a little differently in design technology, just a little different so that when it's being used in the healthcare facility, where we could perhaps narrow the operating temperature, and we could optimize one of the other characteristics of the product.

I hope you've found these remarks helpful. Thank you.

(Applause)

DR. MYERS: Thank you. Do we have any questions for either Dr. Cariski or Mr. Fliss? Yes?

MS. ULLMAN: Ellen Ullman, patient advocate.

I wanted to know with respect to the temperature of the strips, because now a lot of people through managed care are getting their strips delivered through mail order pharmacy, spending at least a week in transportation in dark brown UPS trucks in south Florida, so it's been now exposed to, I don't know, 140 degrees, who knows. Are they still accurate, those strips?

DR. CARISKI: You know, the supply chain and the way strips are transported is taken into account in the way the strips are made and the expiration dates. But what isn't taken into account is the notion that somebody can leave the vial in the baking sun for weeks on end, okay. But those elements of the supply chain are taken into account.

MS. ULLMAN: So up to how many days is that --

DR. CARISKI: I couldn't tell you, because

it's going to vary on temperature and humidity and the product, and there are all kinds of variables. I will say this: in case there's ever any doubt, the best thing to do would be to run a control solution test, and that'll tell you the system's working or not.

MS. ULLMAN: Okay. Thank you.

DR. CARISKI: Sure.

MR. SCHLEIS: Tom Schleis with Octapharma. We're one of the manufacturers of immune globulin that has maltose in it, and we've worked with the FDA very closely to make our customers and our patients aware of this interference. And I published a peer-reviewed article regarding the interference of maltose, icodextrin, galactose, and xylose with blood glucose monitoring systems. And we've hard about maltose and icodextrin. But the other thing I learned when I was doing the research for this article is about galactose and xylose. And galactose and xylose are naturally occurring. You can find them in fruits, vegetables, herbs, and dairy products. Whether or not ingestion of these foods can result in a high enough level to cause the interference, we really don't know. It's

not been studied. But what's more worrisome is that they are heavily marketed in health food stores for a whole wide range of unsubstantiated claims. They're purported to prevent and cure cancer, cure multiple sclerosis, cure bacterial and viral infections, boost the immune system, improve the functioning of the liver and intestinal tract -- and that's the short list.

And they are promoted in very high doses, doses high enough to cause this interference. And they're also specifically noted to be safe for use in diabetic patients, because they don't increase glucose levels -- serum glucose levels. So this is another area where we really don't know if patients are experiencing hypoglycemia and potentially deaths as a result of these two sugars.

And when I saw the number of 13 deaths with maltose and icodextrin-containing solutions, that's 13 deaths that we know of. And as was mentioned, many of these -- more of these cases of adverse incidents are unreported, so that number is, I think we can safely say, is quite a bit higher.

And I have a problem just looking at this as a statistic. To me, this is 13 deaths. This is 13 families that lost a loved one because of an interference that was completely avoidable. And I really have to ask the question, why do we allow these glucometers to still remain on the market? These patients did not die from maltose. They did not die from icodextrin. They died because the glucometer gave a false reading. And we have safe alternatives. We continue to do as much education as possible, but I guarantee there will be more deaths, and I have a real problem with that.

DR. HARPER: Well, we definitely appreciate your comments. And I agree that we may not know the extent of the problem, so any information that you may have discovered about some other potential issues, we would be happy to hear.

We are talking to manufacturers and others to try and figure out what is the best way to move forward to be sure that products are available for people who are safe. And so any input we get on that is actually very much appreciated.

MR. SCHLEIS: Thank you.

DR. MYERS: Before we take the next question, we're scheduled now for our open panel discussion. So I just want to indicate that we're open for any questions to any of the panel members at this time.

MS. SKEENS: Hi. Lisa Skeens, Baxter Healthcare, and I appreciate the comments from the person before me and agree with that.

I want to now move on to what can we to further mitigate the interference issue, whether it's icodextrin or other drug products that are now. For Extraneal or icodextrin, we have a global risk management program that is in place to ensure that we are educating clinicians and our patients about this risk. And we've done verification of that training. We also provide patient med guides; we provide hospital admission kits to educate them about how to advocate for themselves when they are going into a hospital; and this will shortly be approved as a REMS(ph) through Cedar.

And so part of the question is, if you are a

manufacturer who has an interfering enzyme and provides falsely elevated readings when used with something like icodextrin, you know, what are the device companies to mitigate the risk, to educate their end users, such as hospitals, where we're seeing the biggest issue? It's actually trained healthcare professionals where we're seeing the significant issues around the world.

So what can device manufacturers do, or what can FDA do to require risk management programs of device manufacturers to further reduce the risk, to educate clinicians about the risk? And it seems like there should be parity between what we're doing on the drug side and what we're doing on the device side. Other things I know, other ministries of health, have done are required stickers in the hospital. So actually a sticker on the point of care monitors that reminds the nurse that if this person is on peritoneal dialysis, that they should not be using the point or care monitors, that they should be using the analytical hospital labs.

So are there other -- I want to throw out

those ideas and really ask both industry and FDA to comment on that. Thank you.

DR. MYERS: Thank you.

DR. HARPER: I'll just go briefly, Mike, first, because our recommendations are in the public health notification. And just to reiterate, our first recommendation is that health care facilities avoid using these particular test strips in their facility at all, and if they do use them, we did recommend a series of steps that we believe would be risk mitigation steps to help increase awareness of this.

So we have attempted to do that. The question does remain open on how effective has that been. I don't think we have enough time yet to understand how effective that communication has been yet, but we are following it.

MR. FLISS: On behalf of industry, it is our goal to provide safe, reliable, and accurate product, and we think that we're doing so. When we become aware of an issue, we have -- the best solution is to do something through design. But you also, if you're unable to do that for a period of time, you try to

address it through labeling, through an education campaign.

In the case of maltose, you're seeing the companies get the message out over the course of several years, and keep having the message out. And then through design, in addition to having the stickers on the device, some of the meters now have been changed so that there is a software prompt that comes up for the operator to make sure to ask themselves, "Is this a patient who might be on therapy that would raise their maltose level?" So that's a design solution for the short term.

MR. CEMBROWSKI: Cembrowski, University of Alberta. A point for Dr. Cariski. You could get out of your box a little bit. There are different boxes for measuring blood glucose. Probably the best one to compare to any point of care method is the blood gas glucose. It gives you whole blood. It's as good as -- - as good as or better than a laboratory method, and you can use it in your evaluations or even sort of post-marketing, and this would be a good idea for FDA. There are a lot of patients in ICUs who have their

bloods drawn both with capillary specimens that are run on point of care instruments and then they also have their arterial bloods run on blood gas glucoses. And it's amazing what you can find. And I think you could be able to discover many, many interesting outliers that you could use to develop better systems.

DR. CARISKI: Well, thank you for your suggestion. I know that people from R&D are here listening to the meeting. I know that we use as our reference the YSI, which is supposed to be a, you know, reasonably accurate instrument and measures capillary samples so that we're comparing, you know, apples to apples rather than, you know, a capillary to arterial. But thank you for your point.

MS. SOLDO: Hi. Monnett Soldo again from OptiScan.

You know, we in our company, use a different technology, so perhaps it's not a fair question, but I'm going to ask anyway. Where we put the bar for our own performance is that we meet the ISO standard even in the presence of all the interference that one would find in the ICU. And so I'm a little bit

confused in this discussion how we're distinguishing between, there's the ISO standard and then there's interferences, and we're going to just basically try to label against them, which obviously has effectiveness problems; right?

So I would challenge the other industry representatives on two counts. Number one, can we meet the ISO standard as currently stated in an ICU setting, including all the interferences that are currently present in that environment, number one? And number two, is it possible to introduce, using maybe this dynamic electrochemistry or whatever was mentioned before, some sort of a no-read criteria to detect outliers and report them as outliers, and give no reading whatsoever rather than a fundamentally wrong reading?

It would seem to me that the technology possible today ought to support both of those, so I'd like to hear your thoughts.

DR. CARISKI: Well, I've looked around and nobody else looks like they're prepared to answer, so I'll do the best that I can.

First of all, I think this meeting has -- one thing it's done has clearly distinguished between meters for consumers and meters for hospitals. And the ISO standard, by its very statement, addresses consumer meters. It has nothing to do with hospital meters, and by default, it came to be used for meters that are used in hospitals. And as has been noted, CLSI is working on POCT(12) to address standards for hospital meters, and that's a separate issue.

For consumer meters, I tried to point out how difficult it is for manufacturers with current technologies to eliminate all interferences. And when you start stacking interferences, it becomes a real problem. So a meter might do okay with the crit, and then you add some uric acid, and it's okay. But then when you start adding, you know, acetaminophen and the patient has high lipids and a high bilirubin, and -- which, you know, some people do when they're outpatients -- it becomes really problematic.

I don't know enough about the technology that Dr. Ginsberg mentioned to say to what extent it can successfully address these things, but also there

may be patents involved that prevent other manufacturers from using it. All I can say is that every manufacturer, I think it's fair to say, wants to make the best, most accurate product that they can that the patient can use, and we're all working towards that end.

DR. HARPER: I would actually like to press on this issue a little bit more, because I think this is something we're interested in. Not -- not simply for the current ISO standard or some other standard, but if there were to be standards developed, no matter whether it's a standard for lay use or a standard for hospital use, do you believe -- anyone on the panel -- do you believe that interference should be a part of this concept of total system accuracy, or a total allowable error concept? Because, you know, what is encompassed in that total allowable error in terms of the intended use of the device?

So I'd like to hear from you all about your opinions on the inclusion of interference in that.

MR. FLISS: Well, the challenge is when you're trying to stack potential errors to evaluate

what their total might be. What we do is, before we market any new product, we bring our product out to the marketplace, to hospitals, physician offices, and we conduct method comparison studies where we gather samples from patients, measure them on the investigative device as well as on our reference method, and we're also recording what concentrations of potential interferences are in that sample to see, if we do find that there's a flyer, was it caused by the presence of one of the compounds that we suspect is an interferent for that particular assay?

DR. HARPER: So do you think interference should be included in the total allowable error? Is that what you're saying?

MR. FLISS: Well, actually I think you find that by looking at the regression equation that comes out of the method comparison study. Because the sample isn't contrived or controlled. It in theory has whatever interfering compounds the patient is carrying with them or her at that given time.

DR. HARPER: Yeah. I mean, I think we're not aware, and perhaps you have this data and we

haven't seen it or aren't aware -- I don't think we're aware at FDA when the data we get, that it actually would, you know, be a broad enough selection of patients to actually get an idea of the range of interferences that are possible. And that's especially true for, I'd say, hospital use and the types of sick patients. We already talked about that this morning. But even in the lay environment, the sort of ranges of interferences you'd see. So we are interested in feedback from people, you know, to the docket or any other way on how interferences can be assessed, and also which products -- and we're working on this as well with some other groups -- which products are actually most important to look at where you might be more likely to see it more often than sporadically. Because if you have a sporadic interference that's significant, you might want to look at that. If you have a constant sort of interference that's likely, you also want to make sure you look at that. And I don't know how much current study designs actually evaluate that.

MR. ERVIN: I would like to suggest that

interferences usually aren't sporadic; that an interference is going to be a relatively systemic effect. Witness ascorbate as an example.

And another thing I would like to suggest is that rather than including it, necessarily, in total error, eliminate them. I think we should be trying to develop technologies -- and there are pieces of them out there -- that manufacturers can use to ultimately eliminate the effect of interferences. Maltose is a good example. There are other enzymes available. You can get away from it, but -- from a manufacturer's perspective, that's not an easy thing to do, but it's really where we should be headed.

DR. HARPER: Yeah, I mean by sporadic, I had meant, you know, people -- some people are on a drug and some people aren't. Some people may be on acetaminophen, some people may not.

But that's an interesting point of view, I think, is to eliminate it. And I think we would all like to see that happen.

But another question I had relative to this question of interferences is how best to communicate

this information to lay users? I tell you one thing we struggle with sometimes is if there is a certain level of interference in a laboratory or a hospital environment, we may be able to label against it. Sometimes more effectively than others, obviously, with the PQQ issue shows that sometimes not -- not terribly effective. But in the lay population, you may have interferences that are not identifiable by the patient. So they may not know that they have a high level of triglycerides; they may not know that they have some sort of high level of an endogenous compound. They may be able to avoid acetaminophen if they read it. But so any suggestions, as well, on how best to communicate information about interferences to the lay public would also be helpful.

MR. FLISS: We're also interested in pursuing that line of thought. We were recently considering ascorbic acid as being something of interest that shows up in some of the tests for package inserts. But how would someone from the lay public figure out what that means to them? Do they realize that we're talking about Vitamin C, and how

many glasses of orange juice or how many vitamins they can take before they reach a point that they should be concerned? So making a limitation statement that's very scientific in nature isn't easily used by the lay public. So I was hoping that this would be the type of subject that we could address perhaps within this consensus organization, whether it be CLSI, or if it's a lay product, it'd be the ISO organization.

DR. CARISKI: I want to make one comment, and that is that there has to be some limit to the consideration of interferences in terms of total error, because by definition, they're interfering substances. We don't test every substance known to man. We test only certain substances, because we know they have a propensity to interfere with the technology. And sure, ultimately we would like to have a technology that eliminates any interference whatsoever, but right now, to my knowledge, for the test strips that isn't the case.

One thing that patients can do is they can check their blood sugar against the lab. And there are instructions in all the inserts telling them how

to do that, and it will give them some idea as to how accurate it is for them. So whereas they may not know their lipid level or their -- how much acetaminophen or ascorbic acid they have, they can see whether they're close to the lab or not.

DR. HARPER: I will comment that at FDA, we tend to believe that all sources of error should be included in an allowable error measurement. So we would be interested in actually identifying the interferences and being sure that a particular patient is likely to be within that error range, even with the presence of interferences.

MS. PINKOS: Arleen Pinkos, FDA. My expectations for this session were to hear the barriers, and for the barriers to be described specifically enough that we could start working on the solutions towards them. We've already heard that there are some other technologies that can eliminate a lot of the interferences and can get more accurate results. Can we ever get closer to the performance that we need in terms of eliminating all the interferences and physiological limitations with this

technology, or do we really need to be looking at a different technology? Is it achievable or not, and if it really is, I've heard cost is a factor. But I mean, is there anything that's specific that anybody can do to help remove those limitations?

MR. ERVIN: Well, I think that there are at least two companies out there that are demonstrating that you can do this, and publishing a lot on the topic. So like with all things, sometimes it takes considerable energy to move a huge, massive business. I mean you're talking here multiple billions of tests and lots of different products. The companies that can come up with a innovative technology that has a proprietary position and can go and develop tools to do this, they're working on it. They're doing it. Other companies have a lot invested in where they're at at the moment, and they have to move forward, too. But it's a little bit more difficult for them.

MS. PINKOS: Just a follow-up. Are the companies that you're referring to things like HemoCue and i-STAT, or are there a simple, inexpensive device that people could use at home that are starting to

overcome the physiological limitations and interferences? I guess my question is, is it possible for these home use types of meters -- the pure, simple point of care analyzers that we're used to thinking of at the home use -- are we ever going to, with that technology, be able to overcome the limitations, or do we really -- do we need a new technology like some of these other devices and the blood gas analyzer?

MR. ERVIN: I don't know if we can get this technology into the home use situation. Sitting behind you is Jeff Dubois, who probably could address that more easily. But I don't see any long-term barrier to it. As with most things, you -- once you scale up things and you get your technology made rock-solid and less costly to manufacture, yeah, you can spread it out into other environments. And I think that's possible with some of these products. In fact, I know some of them are actually targeting the home use market now. So it's possible.

DR. KLONOFF: David Klonoff, Mills-Peninsula Health Services. One suggestion was made that if there's interference, that the meter should provide no

information, should just be blocked. That ties in with one of the accuracy topics, that we were trying to figure out how to deal with outliers. But I didn't hear anybody this morning talk about how to deal with non-readings, what happens if there's no reading at all? That should be a form of an outlier. I think that should be taken into account.

MS. BOWMAN: Cynthia Bowman, Long Island Jewish Medical Center. I know this is a conference concerning glucometers, but I'm just going to say as far as interference goes, it's not so easy in the central lab on the main analyzers, too. And if any of you have ever tried to confirm your interferences -- you know, confirm the manufacturer's claims? Sometimes you wish you hadn't tried, because it is really confusing. And for somebody who's been on a three-year odyssey right now to try to figure out what's going on because you have -- you put rules in and, you know, middle ware and auto verification and all that sort of stuff. And I'm aware of peers who've gone the same road that I have, that it varies over time. You try to set very conservative limits.

But the other thing, too, and there's literature about one person's lipemia is not the same as another person's lipemia, different particle size -- I suspect the same thing is true for hemolysis and icterus also. And so I can tell you that, you know, I give the device manufacturers their due -- it's not easy. And I think it probably cuts across all sorts of instruments, and the interferent we fear is the one we're not aware of.

DR. SACKS: Sacks from Boston. One of the points that came up in several slides in this last session as one of the limitations of the technology was the time. It's not clear to me why this is necessary. I can understand why patients would want small meters that they can put in their purse or pocket, why they'd want small volume of blood. But why a test needs to be done in ten seconds or five seconds versus 45 seconds is not clear to me if people are only doing it four times a day. Would manufacturers care to respond to that?

DR. CARISKI: I know that -- I don't have the specific data. I know it's been looked at in

terms of patient preferences. And clearly the preference is for shorter, because, you know, it's like anything else. We used to have, you know, a modem that worked at 400 or 1600 or whatever, and now it's -- and you wonder how we ever lived with it before.

And as meters became faster and faster, that was what people expected. I think if a manufacturer were to come out with a meter now that took a minute, let's say, but was more accurate, I'm not sure how many people would -- would embrace it. All I know is the marketing says that most people want faster and, you know, smaller.

MR. FLISS: As people are living more active lives, they desire to be able to discreetly perform their tests. And being able to get a result quickly is attractive to many people.

DR. SACKS: I understand those points, but I remain unconvinced that a 30-second difference in the timing that could significantly improve the accuracy of a meter and avoid hypoglycemia is beneficial to patient care.

DR. CARISKI: The other -- thank you.

DR. WHITE: Neil White, pediatric endocrinologist from St. Louis. I think if -- the patients don't understand the importance or the lack of accuracy that we're talking about here, and if you tried to market a meter now that took a minute next to a meter that took five seconds, they're not going to go for the minute. I can tell you that. They -- every time the meter gets shorter, they say, Oh, I'm going to get this one, because it only takes three seconds or whatever.

Now, if we can really demonstrate an importance in certain patient populations where the accuracy is a key component that they will understand that accuracy as important, then they might be willing to sacrifice the speed for the accuracy if there was a reason that they needed -- that we can convince them that they needed that accuracy. But at this point in time, they don't know -- they're not convinced of that. They're not even convinced that there's an error at all.

MS. KOLLER: Beth Koller from --

DR. CARISKI: I'd like to add one thing, if I may. I know that for some technologies, more time doesn't make a difference. So it may be a matter of going to a whole new technology. It's not like if you have a strip today and you said, Okay, instead of running the test five seconds, we'll run it 30 seconds. It'll be more accurate. It won't necessarily be true, because the strip may have been optimized for the five-second result, okay. So it's a little complicated, is what I'm trying to say.

MS. KOLLER: Beth Koller from CMS.

I'd like to make an inquiry about what is and is not in the label, and how the label changes. And if you would comment on how this relates to drugs, where I think people have a better understanding of what a label means.

It's my understanding that a label can be -- can be more transient for devices than it is for -- for drugs, that there can be -- that manufacturers can insert the label -- can change the labels without agreement with the FDA. And it's also been our experience that it's difficult to actually find the

most current label for a device, and this has some implications for us at our agency.

And in addition, because many devices are approved - - not approved, they're cleared through the 510(k) process, if you could comment on that. And what kinds of drift in the data might occur when you're using not a single predicate, where there can be multiple predicates.

Thank you.

DR. HARPER: Sure. So Beth has pointed out some -- definitely some differences between drug labels and device labels. So for those of you who aren't aware, when a drug is approved, the labeling is available on line, and it actually cannot be modified without interaction with FDA and FDA approval of that modified labeling.

For Class 2 devices like glucose meters, that isn't true. FDA looks at what we would consider draft labeling. And so some of the information on the labeling, or the way that's stated, could possibly be modified by the manufacturer without a submission to FDA or FDA review.

Now, there are some limitations on that. They cannot change claims or add additional claims to the labeling that would be outside of the clearance they received. But they can certainly modify the labeling sufficiently.

And also you're right that there is right now, because of this a little bit, there isn't consistently any sort of resource to obtain device labeling. There's definitely discussions, especially for over-the-counter products, of trying to get that to happen, but those discussions are in the early stages. And it's difficult where there is a draft labeling involved, in terms of figuring out how to keep that current. So that is definitely one of the challenges there.

In terms of the potential problem with performance drift because of the nature of the 510(k) program, it certainly is possible. For those of you who are less familiar, FDA has -- and for Class 2 devices, it's actually a performance between devices. So if one device compares to another device, and that device compares to another device, and that compares

to another device, there is a possibility sometimes for performance drift from the original device. And this sometimes creates a situation where devices aren't comparable.

We try to minimize that as much as we can. I think it's very clear that these devices aren't always completely comparable to each other. But at least for blood glucose meters, we do look at them in relationship to a reference method that's -- we try to keep a little bit more traceable. So we try to minimize that a little bit by comparing the performance to the reference, rather than to another blood glucose meter. It may not be ideal, but it certainly is something to keep in mind.

MR. FLISS: If I may add a remark? There is a certain limitation to that freedom to change labeling, like Dr. Harper mentioned. We're not allowed to expand the intended use or indication for use of the product, and we're not allowed to remove a warning, precaution, or limitation without seeking concurrence from the FDA before that labeling is changed.

DR. MYERS: Yeah, and I have one other comment.

With the 510(k), using the predicate device comparison, what would it take to change that to do it -- make it a standards-based comparison, rather than a predicate comparison?

DR. HARPER: We do have, or have exercised, some leeway there, in that in 2003, FDA actually did recognize the ISO standard. And that standard was actually better than what we had been using before. And so there are devices that are not cleared because they don't meet that standard.

So there is some way for us to do that, and certainly right now many of you may be aware that the Center for Devices is actually looking into the 510(k) program in general. And part of that evaluation is looking at the scientific standards and where they should be looked at or evaluated. So I think if we have, especially clinical and scientific reasons behind meeting certain types of data, I think that where we have justification, we could do it for patient safety.

MR. NEUMANN: Glenn Neumann, New World Regulatory Solutions.

Something that struck me early on today was how much the user can impact the accuracy of the test. And it seems to me we have -- we know about human factors. I don't think we know enough. We have flex study menu; I don't think it's big enough. So perhaps if we could put more effort into preventing user error through flex studies and better technology, even -- if we could take or eliminate or reduce, seriously reduce, the user error, we could take a big bite out of this total error, it would seem to me.

So for example, I've typed how many millions of words into my computer in my lifetime. I still make typos. I still make the same typos over and over. Microsoft Word is smart enough to correct some of them for me. But if we could do something with these meters that would see and detect user error, and if we do more flex studies to really know what goes on, you now, maybe what Courtney needs is a million data points, a thousand users doing a thousand repetitions. The young man who was here did over

100,000. He got a 76 and a 210 on the same day. There's got to be something going on there. So I think if we focus on that, that's something we can do right now. It could help out.

MR. FLISS: Thank you. I know you're hearing a lot about standards today, but I've got a couple more to share with you.

IEC has published a document that has the number 62366, and it describes a usability engineering program that a company might consider while designing a product. Another standard is ISO 14971, which is risk management for medical devices. And there was an annex added to that a couple years ago to address in vitro diagnostic products.

So the designers of blood glucose meters are aware of these standards and have adapted our development, manufacturing, verification and validation activities to comply with those standards. So an example of how a system might be changed because of usability engineering -- many devices now have come out with under-dose detection. So although the size of the sample is much smaller than it used to be, but

still if the user places a sample on the strip that is too small to conduct the test, the meter now is smart enough to notice that and default to an error message rather than giving a biased test result.

DR. SCOTT: Mitch Scott from St. Louis.

One thing I'm hearing now that I think everyone is buying into is that we need to have different performance criteria in different settings. The three of you from industry just said you support that.

The issue today, though, is that it doesn't exist. The technology, the interferences, is pretty much spread across home use and hospital use meters; correct? Okay? That leaves you to a hypothetical question that I think gets to the crux of the matter, and I'd like to hear your opinions on the answer. You have two choices. You tell those of us in a hospital, "Keep using what you're using. We'll get you a better meter in three or five years." The other alternative is stop using what you're using, and go use a blood gas analyzer or an i-STAT.

That's the two answers you've got, if

everyone's in agreement with this, and I --

MR. ERVIN: In the hospital.

DR. SCOTT: -- in the hospital. In the hospital.

DR. CARISKI: Isn't that a question the hospital can answer -- isn't that a question the hospital can answer for itself?

DR. SCOTT: You're absolutely correct. But we're going to have a lot of pressure on cost.

UNIDENTIFIED: He would answer in a different way than his hospital.

(Laughter)

DR. SCOTT: Yes. If I had my 'druthers, we would switch to a more accurate method, but I think we'll run into issues with cost.

MR. FLISS: I think your question is complicated for us to answer, because we haven't really yet considered how accurate, precise a device needs to be in order to address your need to implement a tight glycemic control, so -- and our devices aren't on label indicated for tight glycemic control programs. They're monitoring devices. They were

created when we thought that our self-testers were interested in being between 80 and 180, and I believe it was mentioned earlier today that the hospital protocol has historically been to try to keep the patient below 180 milligrams per deciliter. So that's what we're providing today.

DR. SCOTT: Okay. So, does anybody want to tell me what to do?

(Laughter)

MR. WHITE: I have a question which is a little bit off the subject, and I'm sorry. I wanted to ask somebody in response to something that Mr. Ervin said, and I don't know if you'll be able to answer this or not, but maybe somebody here in the front row can help me answer this. You talked about the po2 and the difference between arterial and venous blood. When we are doing physiologic studies in a clinical research center environment, we are often trying to keep blood glucoses at a steady level, such as by a clamp. We often arterialize the blood because we think that the venous blood is different than the arterial blood. Is that an analytic difference, or is

that a physiologic difference of actual blood sugar?

MR. ERVIN: And the process for arterializing involves?

MR. WHITE: Warming -- warming the hand.

MR. ERVIN: Okay. You're going to have the temperature effect on the equilibrium. I'm going to say that that's probably an analytical effect.

UNIDENTIFIED: Think so? (Inaudible, off mike) It's different than what I've always been taught, yeah.

DR. GINSBERG: When you arterialize the blood, what you basically do is you create AV shots, so that you basically bypass the capillaries, and the arterial blood is now picked up in the veins. And when you do that, you do two things. One, you do increase the oxygen. Well, they don't think you increase the oxygen in the arterialized blood much above 80 or 85 milligrams -- milliliters of mercury. So you're not going to create a major oxygen problem, I don't think. I've never checked it, but I don't think you're going to create that.

What you are going to do, though, is there's

another factor. And that is, the blood glucose is not the same in the arteries, the capillaries, and the veins. And part of that is because there's some glucose extracted by the muscles and other things as you go along there. And when you arterialize it -- and for that reason, arterial and capillary blood are actually pretty close. Venous blood is lower, particularly in a time around a meal. So that if you measure glucose in the veins about -- within an hour to two hours after a meal, it will be one to two millimolar. So 18 to 36 milligrams per deciliter lower than capillary blood, which is very similar to arterializing. But because arterializing capillary is similar, and I don't think there's an oxygen problem on arterialized blood, although I don't know that, it'd be worth checking -- I think that arterialized blood would be very similar to capillary blood in terms of the number you get.

MR. WHITE: (Off mike) What you're saying is what I've always been taught, but you (inaudible).

MR. ERVIN: I probably misunderstood the point you were trying to make. If in fact the oxygen

levels in that arterialized venous blood is in the order of 80 to 90 milligrams -- I'm sorry, millimeters pressure, then that's very -- very, very close to capillary level. So you're not going to see a po2 effect with these po2 sensitive methods in that application.

MR. STR Hi. Anders Strfrom HemoCue Sweden.

A question for FDA. Considering the two sessions today talking about accuracy and precision, combined with this interference discussion we're having, is the result of this that we need to have different requirements for different type of patient settings and so on? So it's not only about home use and hospital use, but it's about neonatals, about different type of ICU settings, about different type of home patients, being that -- using a lot of -- I don't know the English word, but substances you buy in the drugstore and so on?

DR. HARPER: So I think that part of -- part of the reason that we hosted this meeting is actually to hear from this community of people about maybe what steps we might need to take. Because, you know, we

definitely heard the point of view that perhaps there needs to be some enforcement of FDA's point of view about intended use population. So, you know, right now these meters are being used off-label as part of the practice of medicine in hospitals, and we're hearing some people have feedback that perhaps they're not safe.

We've also heard today, though, that some people are using them in their hospitals and need them. So we'll have to create the right balance, and we're interested from hearing all stakeholders on that. But that is exactly the question that we're trying to address, is how do we make these safe for all patients on which they're used?

DR. BRETON: It just tends to be just home use versus hospital use, and that means maybe not the only two options.

DR. HARPER: Yeah, I mean, quite frankly, even though these meters come in for over-the-counter clearance right now, when they are claiming healthcare provider use in populations such as neonates, we already actually ask special questions for those. So

I think it wouldn't be outside of our realm of comfort to be sure that if there are populations that need to be addressed, we would want to address them.

MR. DUBOIS: Jeff DuBois. I'm with Nova Biomedical, Waltham, Massachusetts. And Ken used my name, and indicated that there may be some technology. So in answer to Arleen's question, there is some technology that does address the issue of accuracy and precision. Barry referred to it this morning.

But that's not why I'm up here. Gary has been involved with a program through NKDEP where we looked at an analyte that's problematic in the laboratory. That's creatinine. And there's an initiative that's global to standardize creatinine measurements so that we can report an EGFR and properly assess the patient's glamerial filtration rates and assess their stage in chronic kidney disease.

What we haven't done with glucose, and what we need to do, and having been responsible for the Area Committee for Point of Care Testing of CLSI, is to have an initiative where we standardize glucose

testing. There was some reference to IDMS traceability. And there are informed or notarized bodies in Europe that take some different methods than we do here in the United States, and there's traceability.

So what we need to move toward in the United States and overseas is harmonization of this approach with glucose. And once we begin to do that -- industry, practitioners, and the regulatory bodies -- then we will improve the accuracy and precision of devices.

DR. MYERS: Yeah. A comment on that. One of the things that we've looked at at CDC in the past is looking at that very issue of having a standardization program for blood glucose meters. And the issue is -- one of the challenges that we have is that unlike for creatinine, we don't have good reference materials that really simulate what the blood glucose meters are actually measuring.

And that's part of the challenge that we're facing: how do you come up with a whole blood glucose reference material that's easily stable, that can be

used to establish traceability? So there are challenges involved in coming up with a program that standardizes blood glucose meters. Not laboratory instruments, but blood glucose meters.

MR. DUBOIS: But Gary, there are issues with the central lab method. I attended a chronic disease conference, and David was the speaker there, along with a fellow from UCLA, Davidson. And their justification for using hemoglobin A1C for diagnosis of diabetes is there's too much variability in plasma glucose from central lab analyzers. So we've got a problem, and we need to begin to work at it as a community. So I don't see glucose apart from central lab or apart from self-monitoring, or at the bedside. Glucose is glucose, and we really need to give accurate and precise glucose measurements to our clinicians so they can make appropriate decisions about the use of a lethal drug, insulin.

MS. PINKOS: Hi. Arleen Pinkos, FDA. I have two questions, and I'd really appreciate hearing an answer to both of them.

The first is, we've already heard this

morning that there are some technologies in point of care meters that are meeting a much stricter performance criteria, or performance cutoff, closer to ten percent. What can be done to get some of the other manufacturers to pull their quality up? Is that something that can be done when incentives might be provided?

And secondly, the way FDA operates, once a product is on the market, and maybe in five years or seven years, it still doesn't meet the current requirement, should there be something in place, like a sunset law, that says, you know, here's the new performance criteria. If you still haven't met it in five years or whatever that time might be -- is that a reasonable approach? Or should they all be left on the market indefinitely?

MR. ERVIN: I'll try and answer the first question. The second one, I'm probably going to leave to either industry or to Dr. Harper here.

I would be surprised if every manufacturer is not already trying to get to that plus or minus ten percent. I know that there are programs in development

that have that as their target, and so I think it's on the near horizon. And that's about all I could say on that is that there are -- it is an important goal of all of these companies, to get there.

DR. HARPER: I don't know if I can address your second question or not, Arleen.

MS. PINKOS: No, I really wanted to hear from --

DR. HARPER: Right.

MS. PINKOS: -- from Mike and Alan, from industry, like do you think that's a reasonable approach to have some type of sunset law on -- one people. When the performance requirements go up -- let's say it's raised to 15 or 10 percent now, then in another five years or whatever, it's raised again -- what, if any, action do you think should be taken on all those products that were cleared a while ago and they're not anywhere near that -- that criteria? What's -- what's your opinion on that?

MR. FLISS: My sense is that individuals, when they grow accustomed to using a particular test system, it becomes part of their daily live, and they

would like to continue to use that system if they have found that it's safe and effective for managing their particular health. And although the 510(k) may have cleared three years ago, if it's a reliable product for that individual, I think industry would like to continue to provide test strips and controls so that they can continue to operate that meter.

DR. CARISKI: I think also it should be based on risk, as best one can assess it. And for example, you know, it may be that one wants to look at different risk categories. So for example, as Dr. Ginsberg suggested, the standard for someone who's a Type II on orals or diet, and the oral isn't anything that will produce hypoglycemia, the current standard may be okay. And if it turns out that the meters meeting the tighter standard are a lot more expensive in terms of the meter itself and/or the test strips, it may not be appropriate to require people who don't need that kind of accuracy to get a more accurate meter. But it may be appropriate to label the -- but to, you know, make that distinction in the labeling as to who the intended user is.

Again, it's a question of health risk. I mean, I don't think anybody today would remove all the meters from the market on the grounds that they're not as accurate as we would like. They're certainly doing an adequate job. You know, it could be better, but they're certainly better than nothing. So to just like say, you know, in X number of years if the standard is stricter, we should remove all the old meters, I'm not so sure I would agree with that. Partly for the economic reasons and, as Mike said, when people get accustomed to --

MS. PINKOS: That is what they do in Europe, though, right? With the IVD directive you have a certain grace period and then if you don't need it any more, you're -- you have to come off the market?

DR. CARISKI: That's true.

MS. PINKOS: Do I understand?

DR. CARISKI: It's not like it's not been tried. You're just asking our opinion. I'm giving you our opinion. You didn't ask what's done worldwide. Thank you.

DR. MYERS: This will have to be our last

question.

MR. WHITE: Yeah. I'm sorry to be up here again. I guess I just like to hear myself speak. But I just wanted to make a comment that I don't think had been made earlier until just a minute ago, when Dr. Harper made another setting in which we really have to be very careful is our very vulnerable population of the neonatal intensive care unit, where sugars tend to run on the low side anyway, so the accuracy is poor. And they have all -- many of the different interference factors, such as different hemoglobins and many medications and low oxygens and high PC02's. And I think we have to really think of that as a population which we don't know if we're getting any good numbers or not.

DR. MYERS: Well, that concludes our afternoon session. I'm going to turn the microphone over to Dr. Harper for closing remarks. Before I do, I want to thank all of our afternoon speakers.

(Applause)

DR. HARPER: So I really want to thank everyone for a wonderful day. As I said in the

opening of this particular session, I'm really happy with the -- the information that I've heard with the discussions that we've had so far, and I'm really looking forward to tomorrow.

So as a reminder, we start again tomorrow at 9:00 a.m. And everyone who's here today, you still will need to sign in again tomorrow, as you did this morning.

We have a really exciting day tomorrow. While today we talked about accuracy standards in general and some of the issues with meter performances and interferences, tomorrow morning we're going to focus on the issue of tight glycemic control in hospitals and the advantages and disadvantages of that, and have probably a little bit more discussion on some of the requirements that might be required for that, or the advantages and disadvantages, and hopefully with an emphasis on patient safety.

And then also we are going to hear from a patient representative on the issues that are important to patients when they choose and use blood glucose meters, and also from a point of care risk

manager about the issues that are -- that are important to them.

Opening the day, we'll actually hear about liability issues, potential liability issues on the use of blood glucose meters.

So I hope that you're as excited about tomorrow as I am, and I will adjourn the meeting for the day, and see you tomorrow.

Thank you.

CERTIFICATE OF COURT REPORTER

I, NATASHA KORNILOVA, the officer before whom the foregoing meeting was taken, do hereby certify that the statements was taken by me in audio recording and thereafter reduced to typewriting under my direction; that said transcript is a true record of the recording taken by me; that I am neither counsel for, related to, nor employed by any of the parties to the action in which this transcript was taken; and, further, that I am not a relative or employee of any counsel or attorney employed by the parties hereto, nor financially or otherwise interested in the outcome of this action.

NATASHA KORNILOVA
NOTARY/COURT REPORTER