FOOD AND DRUG ADMINISTRATION

CARDIAC TROPONIN ASSAYS PUBLIC WORKSHOP

Tuesday, November 28, 2017

8:51 a.m.

FDA White Oaks Campus

Great Room, Building 31

10903 New Hampshire Avenue

Silver Spring, MD 20993

Reported by: Natalia Thomas

A P P E A R A N C E S

Fred Apple, Ph.D.
Stayce Beck, Ph.D. (Moderator)
Jeff Bishop, Ph.D.
Paula Caposino, Ph.D.
Anna Marie Chang, M.D.
Robert Christenson, Ph.D.
Christopher deFilippi, Ph.D.
Rakesh Engineer, M.D.
Dina Greene, Ph.D.
Alberto Gutierrez, Ph.D.
Allan Jaffe, M.D.
Kellie Kelm, Ph.D. (Moderator)
Juliane Lessard, Ph.D.
Courtney Lias, Ph.D. (Moderator)
James McCord, M.D.
Richard Nowak, M.C.
Norberto Pantoja-Galicia, Ph.D.
Frank Peacock, M.D.
Jane Phillips, Ph.D.
Ian Pilcher (Moderator)

A P P E A R A N C E S (continued)

Karen Richards
Amy Saenger, Ph.D.
Rick San George, Ph.D.
Yader Sandoval, M.D.
Brittany Schuck, Ph.D. (Moderator)
Zivjema Vucentic, M.D., Ph.D.
Kerry Welsh, M.D., Ph.D.
Jacqueline Wieneke, M.D.

C O N T E N T S

P R O C E E D I N G S

DR. CAPOSINO: Good morning. I think we have everybody, well not everybody but at least our first panel. So we are going to go ahead and start.

All right. I'm going to try to gain some time here.

Good morning. Welcome to White Oak and thank you for being here. Thank you for traveling and waiting in our security lines. We do appreciate you making the travels out here to be with us here today.

My name is Paula Caposino, and I'm the Branch Chief for Cardio-Renal Diagnostic Devices.

During this workshop we would like to discuss questions we all face about troponin devices. We will share some of our experiences and observations with these devices and the goal today is to get feedback from you and to open the lines of communications between all of the stakeholders. This is helpful because sometimes we hear information about FDA's expectations or restrictions on these devices that are not true.

For example there is an idea that we're not

open to high sensitivity troponin devices or that we mandate clinical cut-offs that sponsors are allowed to use. And again these are not true.

We want to work together with all stakeholders so that we can increase the availability of troponin devices that work well and are innovative.

So hopefully you all got the general information. The bathrooms are behind the kiosks and the workshop is being webcast. And an archived webcast link will be available for future viewing. Transcripts from the workshop will be available in approximately 45 days.

During this workshop we will hold a series of five panel discussions. FDA will open each panel with a brief presentation and provide panel discussion topics. Each panel session has time allotted for questions from anyone attending the workshop.

The workshop will also have a public comment session and everyone here is welcome to speak during the open comment session. If you would like to speak during this session, please add your name to the list at the registration desk. And plan to be in the room a

little early in case the workshop is running a little ahead of schedule. But maybe that won't happen.

FDA invited registered participants with relevant experience to participate in the panel discussions. We also invited participants from industry and to do so we have reached out to associations representing device manufacturers to identify interested participants. There will be two members of FDA participating in each panel, a third FDA member will moderate each session.

And again all workshop participants are welcome to speak during the open comment period.

At this time I would like to welcome the panelists for the first session to come up and take their seats.

The participants in this first session will discuss Cut-Off Determination Studies and Reference Interval Studies.

Each troponin device includes information about the reference interval. Since troponin devices are not standardized or harmonized test results are not interchangeable from one device to another. This means

that each device has its own reference interval information.

Now how are these used? The reference interval studies for troponin assays take an important role because current clinical guidelines describe that the clinical cut-off for troponin devices should be based on the 99th percentile upper reference limit of a healthy population.

Many sponsors choose the 99th percentile as the clinical cut-off for their device. However, some sponsor choose to conduct pilot cut-off determination studies to establish the optimal clinical cut-off for their test. When sponsors choose a different cut-off other than one based on the 99th percentile and their device can measure the 99th percentile we request that the sponsor include clinical performance information using the 99th percentile in the labeling in accordance to the clinical guidelines since it is our understanding that this information is useful for the clinical labs and clinicians.

Since these studies pre-specify the clinical cut-off they are important because the more confident a

sponsor is in the pre-determined cut-off or cut-offs the better chance they have in successfully validating the test.

Device manufacturers take different approaches to designing their reference interval studies. For example there are differences in the inclusion and exclusion criteria used to enroll. Some sponsors use questionnaires, others perform testing to exclude sub-clinical disease. Sponsors may also use different statistical methods to analyze their data. Some may want to claim sex specific upper reference limits, others may choose not to.

Our approach has been very hands-off. We try to make sure that the labeling includes information about the population studied and the statistical methods used.

During this panel we look forward to discussing best practices, lessons learned, and if there is a right approach to designing and performing these studies, we look forward to hearing what that may be.

During this panel we have the following

discussion topics: Discuss best practices for trial design including the subject to enroll, how to analyze the data, what to do with results that are identified as outliers, and we offer the question why do many sponsors choose to perform large reference interval studies, and what information about these studies would be helpful to clinicians and laboratorians.

I would like to open the discussion and ask the panelists to introduce themselves.

Stayce Beck will moderate the session.

Thank you all again for taking the time to be here today.

Cut-off Determination/Reference Interval Studies

DR. APPLE: Good morning. So my name is Fred Apple. I'm from Minneapolis, Minnesota. I am Director of Clinical Laboratories at Hennepin County Medical Center as well as I'm a professor at the University of Minnesota in the Department of Laboratory Medicine. And I think I submitted some conflicts of interests or disclosures, so you can look at it. Those are on the website. Thank you.

DR. BISHOP: Hi, my name is Jeff Bishop. I'm

the head of Diagnostics and R&D at Singulex, Incorporated in Alameda, California.

DR. CHRISTENSON: My name is Rob Christenson. I'm a professor of pathology at the University of Maryland, School of Medicine right up in Baltimore. Clinically I direct the Core Laboratories at the University of Maryland Medical Center and am also Medical Director of Point of Care Testing at the University of Maryland.

DR. GREENE: Hi, I'm Dina Greene. I'm an Assistant Professor at the University of Washington where I direct the Core Laboratory.

DR. SAENGER: Good morning, I'm Amy Saenger. I'm Associate Professor of Lab Medicine and Pathology at the University of Minnesota. I'm a Clinical Lab Director and I director our clinical trials research lab.

DR. PENTOJA-GALICIA: Good morning. I'm Norberto Pantoja-Galicia, I'm a mathematical statistician at the Division of Biostatistics in CDRH, FDA.

DR. WIENEKE: Hi, good morning I'm Jacqueline

Wieneke. I came to the FDA in 2011 from clinical practice with a background in internal medicine and anatomic and clinical pathology. I'm currently a Medical Officer in the Division of Toxicology and my responsibility on the review team is as the medical clinical consultant.

DR. BECK: Okay. Good morning. My name is Stayce Beck. And I'm in the Division of Chemistry and Toxicology at FDA.

So we are excited to get this panel started. I won't call on specific people unless conversation is lagging. But we have several discussion topics that are on the board and the first one is really what are some of the best practices for study designs to determine reference intervals for troponin assays. So let's really start with who should be enrolled in the reference interval studies.

DR. APPLE: So one of the things I do do is for the last 30 years I've run a clinical trials study laboratory. We are called the Cardiac Biomarkers Trials Lab. So I have been fortunate to be involved probably with almost every manufacturer in the room

doing either a point of care study or a central lab study, going back to the days of CKMB. And to hear Paula's comments I'm going to take home a message from that because I understand the FDA's role is regulatory. So I'm going to kind of throw the onus maybe on manufacturers because if the FDA can't dictate and tell you, but they just advise, maybe all the manufacturers should get together in a group and decide on what rules they should follow for how many patients they enroll in a reference range study. I mean I chair the IFCC Task Force that has published a document that says a minimum of 300 men, a minimum of 300 women. Yet we all do see many, many different individuals up to thousands of patients. So it was never clear, we could have this discussion, where that number comes from. Is it FDA driven? Is it manufacturer driven?

But I will put - I'll just throw that out as a starting point and maybe we can develop this discussion. If every manufacturer had an annual meeting and say we're going to design, whether this, we're talking here about reference range interval studies, why don't we decide on a set amount, how they are

enrolled. We've decided and Rob and I sit on an AACC Academy, and IFCC group that has just had a submission and I'm sure Rob will have a lot to say about that. The IFCC started off saying we had a number quality, we want a health questionnaire and we have alluded to using surrogate biomarkers to help pine down abnormality whether it is estimated GFR or NT-proBNP or hemoglobin A1C. So I'll just stab, I'm sure I'll come back to the discussion. A lot of people have things to say. And I think that is something we should talk about.

And the other thing I'll comment on is we had a recent paper, we do some work on the new Gen 5 Roche assay and what we did in our paper published in Clinical Biochemistry we looked at the concept that Paula brought up is how do we analyze reference ranges. And in that we looked at the three major methods. We looked at the robusts. We looked at the Harrell-Davis. And we looked at the non-parametric. And if you look closely the numbers change. The numbers change by one or two. And since I'm a member of the Global Task Force the 99th percentile is not going to go away in my

opinion for a while. So how you uniformly decide one company to the next which metric you use and then which outlier process you use. Do you use the two-key or do you use the Dixon whatever, I'm not a statistician. I think we have one on our board. I'm throwing out a lot of ideas to get the conversation going.

But again I'm going to throw it out maybe the manufacturers should come up with a plan, submit it to the FDA and if they said it looks good then you can all uniformly perform your studies.

And I can talk all day as you know. So I'll just stop and let the group move on.

DR. BECK: So thank you. Let's talk about you brought up the idea of study numbers. So I think sometimes we see that these studies do get larger and larger and thousands of people and sometimes that might be appropriate and sometimes it might not be. So let's go ahead and throw out sort of what are people thinking when it should and shouldn't be that large?

DR. CHRISTENSON: Yes, so I think your first bullet here the subjects to enroll and test. I think that should be sort of an all-comers population that

then gets screened. We know that if we use the 300 and 300 especially in the context of troponin where we don't talk about the 97.5th percentile, we talk about the 99th percentile; right. So let's think about the numbers for a moment. If you have 300 men, 300 women and you are trying to get sex specific interval; right, so if you use 300 how many points are you relying on? So one percent of 300 is three points. So it shouldn't be surprising to any of us who are trained statisticians or not that our 95% confidence intervals are very wide. So this I think is the reason, again you don't need a weatherman to tell you which way the wind is blowing, when you've got a 95% -- or a 90% confidence interval, that is what IFCC uses. It is 90% confidence intervals that are so wide that you need larger numbers to make that confidence interval more reasonable.

So 300 and 300 may be a number and maybe it is correct for 97.5th percentile but as laboratorians we like to have at least ten or better 20 values around that cut point so that we can better define exactly what that cut point is. And I think to get that many

you need to have larger n. I mean a thousand will give you what? Ten. So 2,000 probably would be a number that if somebody were to ask me I would say well, 20, why? 20 at least give you some noise around that 99th percentile. So I think that is the reason that folks are using such large data sets.

DR. BECK: So anyone else have any thoughts on sort of what these minimum numbers should be particularly as you mentioned sex specific cut-offs as well as different age groups where it might be impacted?

DR. APPLE: If it is quiet I'll make a comment. So I think -- Amy, go ahead.

DR. SAENGER: No, I was just going to make a comment on the age specific cut-offs. I know Rob mentioned that it should be kind of an all-comers population but I think that is different when you talk about an all-comer emergency room population versus what we are talking about is an all-comer normal healthy population which is going to be a different probably age distributed population than those individuals seen in the emergency room. They are kind

of two different studies in my mind. And the 99th percentile should be in more normal, they'd probably be younger individuals. By the time you are 70, 80, I don't even know if biomarkers could help you with defining if you are truly normal without doing imaging which of course is cost prohibitive.

So I think you could look at the distribution of troponin in a typical ER population in terms of age. But in terms of 99th percentile I think it should be more truly healthy normal individuals.

DR. PANTOJA-GALICIA: I want to mention that it has been mentioned that the number of subjects is important and it certainly is especially if we want to obtain confidence intervals. But also I want to bring a very important point that it is not only the number of subjects but also the variability that we observe that drives the precision in the confidence interval. So if we have more variability the precision is going to increase and if you have less variability we're going to have better precision. So that means also shorter confidence intervals.

DR. BISHOP: So I think talking about the

number of subjects that need to be enrolled and the confidence intervals around them, all of those things come as a result of the fundamental issue which is that troponin concentrations are really a continuum and when we are trying to impose a cut-off of what defines an acute event versus not it becomes an issue there. And then also I think another issue is we start to see data coming from more and more high sensitivity assays. We see that the traditional definition of normal is not necessarily healthy.

And so I think the reason that many sponsors are enrolling more and more people is because they see these wide confidence intervals and the reason they see the wide confidence intervals has to do with the inclusion and the exclusion criteria of the population. And just because someone is self-declared healthy doesn't mean that there isn’t some sub-clinical cause of a troponin elevation in there. And all of those things I think confound the issue. So you can measure 300 or you can measure 3,000 but if you have a different inclusion criteria from sponsor to sponsor none of that is actually going to help.

DR. APPLE: Which raises the point very importantly is that we all know you pick up a package insert they have a 99th percentile. Let's say for the high sensitivity assays whether it is FDA or in the literature. Men are higher than women. Then there is another study done somewhere else the numbers are different. So I think we have to understand that whoever we enroll if we declare how they are determining normality if we use surrogate biomarkers the numbers are going to be different from location to location depending on how the gender, depending how the ethnicity, the race, and the age. I mean who is normal in this room? Does anyone think they can participate in a normal reference range study in this room? Because for example when they had the universal sample bank we did at the AACC in Atlanta, we had a health questionnaire and then we used surrogate biomarkers as I mentioned. And it excludes people. And that was somewhat of an international population. So I think you can't get hung up that your number -- I would say the FDA shouldn't even worry about, the numbers are going to be different from every normal reference range

study. It is unlikely to see exactly the same values that come out of two studies.

So, therefore, I think choosing a number 1,000 or 500 Rob was right, if we chose a 90% confidence interval. If you can do a 95% confidence interval the number goes up statistically like 385. But again if we had some uniformity on all of those issues, how we use surrogate markers and manufacturers have to be open to publishing it in their package inserts. We don't see that.

So we lose confidence as a laboratorian when I don't see exactly how you've defined it. What's the number of men, number of women, what's the reference ranges for them, how many people were excluded from Part A to Part B. We don’t see that. Then you go to the peer reviewed literature and say well, I'd rather use a peer review literature package insert where all that is disclosed than a package insert. But I know too much. The average clinical pathologist that runs a laboratory and Rob can comment on this in the country they are going to take your package insert and that's going to be the gold standard. So the more you can

disclose to the user the better.

DR. BECK: A lot of good ideas have come up that I think need a little more flushing out. So the idea of sub-clinical disease has come up. So how do the panel members recommend testing for sub-clinical disease to really be able to say someone is "normal"?

DR. CHRISTENSON: So as Francis Bacon said so many years ago one half of wisdom is getting the question right. So I guess if the purpose of this is to define a 99th percentile with really good confidence then I think the population probably has to be the all-comers. And the reason I say that is we're comparing a normal healthy population of college students or medical students to a really sick population that is coming in that is symptomatic into an emergency department. I would venture to say that the median age of that would probably be 60 or much higher. So you are comparing two different populations for the same cut point.

So is that really something that evidence based medicine would have us do. I haven't seen anything here about an ROC curve or even should that be

part of this whole discussion. Well, that goes to adjudication and I really look forward to the clinical trials, folks talking about adjudication and how that is best done. What should be the gold standard?

So as Fred brings up it brings up a whole host of questions. But I think if we are going to do a 99th percentile I think we need to do the N that is going to give us whatever the population we decide on has to give us an N that will give us a reliable so that from population to population you don't get different values for different package inserts that our sponsor colleagues get just because that's the randomness of the nature of the thing that we are trying to measure.

DR. PANTOJA-GALICIA: I think you have mentioned something very -- well, the previous two points are very important. I want to bring these two points. The first was uniformity. I think uniformity is important to have consistent or similar estimates because what has been happening is that if the distribution from where we are obtaining our 99th percentile is changing meaning that for example it is tightening, the institution is getting tighter and

tighter then the corresponding 99th percentile is going to continuously be changing. It is going to be moving as the distribution moves, as it tightens. So uniformity is important for that.

And another point that I want to talk about also is that these population is not an intended use population because as it has been mentioned these are the healthy normals. I was reading a paper that Jackie brought to my attention where initially people have the intended use population and then they selected the cut point, the threshold based on that population so that you can get a trade-off between sensitivity and the specificity false positives and true positives based upon the publication. And I understand there are reasons why that has been moved but maybe it is something to consider again and that's all I have.

DR. APPLE: So the concept of normality, we do a lot of reference range studies and we go across and the FDA requests people in their 50s and 60s and 70s. As you get higher in age like you can't find a normal Allan Jaffe everywhere right. So it is hard to get people that are normal as you get over 60. When I hit

60 I had my first measurable troponin value. On the old assays I never had a measurable value above the LoD. Something happens at 60 whether its program cell death or what, all of a sudden I had a measurable value and wow that is interesting. So the point I'm trying to make is if we vet these people out, I don't think you should ever use college students for a normal reference range study as we go around the room and vet people out, are you on statins because you have cardiovascular disease you probably shouldn't be considered normal. If you have a high NT-pro or BNP, if you have a low estimated GFR, so as we vet them out, yes, statistically you lower your values. And that's a good thing because what do we want to do. We want to pick up disease and it is more than diagnosis as we'll hear about it is outcomes management. I mean a diagnosis becomes secondary and I'm not a physician but we want to know how I'm going to manage that patient and how I can improve their outcomes.

So that said there's a lot of value knowing in the age distribution that have been vetted down to -- pine it down from ten to 11, excuse me, from eleven to

ten to eight as an example. I'd rather have that eight because if I show up with a nine I want Frank Peacock to say hey something is going wrong with him. I might have to look at him as an outpatient which is fine. So I like the idea that we will get a tight lower reference range because I don't want to be missed in that noise that could be out there if we enroll everybody without some rules of pining down the number.

DR. SAENGER: I guess I'll just comment on that too. I think I mean if you look at the universal definition of MI you look at all the clinical guidelines they state a healthy normal population not adjusted for you know the age distribution in the ED. And so I think I mean trying to mimic that as best as possible will give us I mean we want the best sensitivity, not have to worry about trade-off. And the reason that we can have the highest level of sensitivity is because the other recommendations have serial testing which increases your specificity. So I think having the two in combination should give you the best clinical performance and outcomes. So if you sacrifice your reference interval setting of a higher

threshold for the 99th percentile then you might have better specificity but you will be as Dr. Apple mentioned probably missing people as well.

DR. APPLE: As you will see in the fourth universal definition that we will hear about later from I'm sure, it is a document that talks about myocardial injury in addition to myocardial infarction. Remember this is a test of injury, not just MI. We lose track of that sometimes. So we want to know if there is an underlying injury. That's part of defining normality. And again I'm going to emphasize if the manufacturers could work together to look at what the experts have proposed in the peer reviewed literature and come up with a number that would be great because then what you bring to the FDA would be uniformly agreeable.

DR. CHRISTENSON: I think what Fred brings up is terribly important and that is the fact that you know all myocardial infarctions have an increase in troponin but not all increases in troponin are myocardial infarction. And so it is myocardial cell injury. So in that way we might want to think about this reference or this normal values so-called or

reference interval as agnostic to which disease that you're looking at whether it be myocardial infarction, pulmonary embolism, renal disease, whether there are other things. So I think again it depends on which question we're asking.

The other thing that I notice you've talked about is outliers and the issue with outliers. And I think we have to decide whether the value we're talking about is an outlier or that the patient that had that value measured is an outlier and should be removed from the population. And we're talking about and so there are a number of statistical tests that can be used. But I think the major thing there would be to determine whether it is the patient or the value that is the outlier. And there are a number of ways to do that but I think that is something we should decide as well.

MS. BECK: Courtney.

DR. LIAS: Hi, I'm Courtney Lias, I am with FDA and I really appreciate the discussion that's going on right now because I think it highlights part of the reason that we wanted to have this particular panel. There are a lot of questions that we have about what

the right population ought to be because we get asked that question a lot.

And I do want to clarify a couple of things that have come up. One, FDA does not currently mandate any particular population for the 99th percentile for the main reason that there doesn't seem to be consensus on what population to use. And two, I don't know that we are necessarily the people that have the expertise or point of view to decide what population ought to be the right population to set a clinical cut-off for troponin.

So you know we are interested in hearing from the clinical community if there is a way to get consensus on either sort of one method of doing it or various methods of doing it that could be transparent to laboratories at the end of the day.

I heard the suggestion that maybe manufacturers ought to make proposals to FDA and FDA ought to decide. It's not my favorite option; I will tell you that. And I think for the reasons I just said that I think it would be more valuable to have some sort of consensus or discussion from the clinical

community on what laboratories, what ER docs need, what is the information that is needed.

So one thing I would like to ask the panel to discuss with that in mind is given that we aren't in a position to decide the right population we have allowed manufacturers to simply describe the population that they used. We haven't mandated anyone do testing to make sure people are healthy. We haven't mandated ages. We haven't done any of that. We simply asked the manufacturer how did you do your study. And tried to make sure that was as transparent as possible to the people using it. So I would be interested to hear the panel discuss a process by which, you know it is okay if we can't do this today, I understand this is not an easy question, a process by which you would suggest we might come to consensus eventually in the clinical community either on one way of doing it like I said, maybe there is one way to do this for all of the types of uses for troponin you mentioned. Or maybe that there are different ways of doing this for different types.

So if anyone has suggestions on sort of next

types of discussions or other discussions that are broader we'd like to hear that. And also who should be involved in such discussions to develop some sort of consensus on this type of question in the clinical community because I think it would really help us and the manufacturer as well.

DR. GREENE: I think I'll take this question first if that is okay. I think there is an abundance of literature that is showing all the different 99th percentile cut-offs that have been determined in different ways. And then also how applying those clinical cut-offs to a clinical population in the ED affects outcome. So I think it is really the integration of those studies to then determine from what has already been done what might be the best approach in defining that population. And I think that the group of experts that need to analyze that data need to be not only the people that participated in those studies but unbiased people from the ED, from cardiology, from laboratory medicine as well as the experts that designed the study and obviously the manufacturers as well.

DR. APPLE: So I'll just add to that before Frank. So Courtney, thank you for your comments. But I think that there are a couple of expert opinion groups out there. My IFCC Task Force and the AACC Academy has an international group of over a dozen people that have been looking at this question. So the question we've come up with is we have a paper in review which I am sure will be accepted at Clinical Chemistry maybe this afternoon, I don't know, any day. And it really clearly spells out, we don't come up with numbers but we're going to have an IFCC meeting. Maybe we have to increase the number from 300 but we have clear guidelines of a broad group of people by ethnicity, gender, age. We have actually designated surrogate biomarker cut-offs we recommend. I think the only thing left for that group to do is to come up with a better number than the 300/300. Maybe we can get a bigger number and in addition we're talking about what statistical method. And I can tell you the ones in our experience for the IFCC, the robust method doesn't work unless you can measure over 50% of people, the numbers fall out. You can't even get a 99th percentile. That

leaves the Harrel-Davis and the non-parametric which there'll be some discussion on that but I think there can be - there is stuff out there that we can have in the literature and it is comprised of ED physicians, cardiologists and laboratorians, so everyone has a seat at the table at this group and it is an international group. So we do get input and so I think we can forward you our final documents when we get them so you can see that. And we tell the manufacturers that when we do studies but no one seems to follow what we suggest. So there is a conundrum here. You don't give advice and then experts give advice and I'd say most of the time neither advice is held. They do -- maybe get other advice. Rob gives them advise maybe different than me, same as Dina, same as Amy. So there's really -- everyone has Rob says what's the question. They hear different questions from us. But I think that is the key. There are some papers out there that will help answer your question.

DR. BECK: Go ahead Frank.

DR. PEACOCK: Frank Peacock, Baylor. So I am an emergency doctor and I have a question for the

panel. Is what is the fixation with the 99th percentile? Tell me the physiological basis for it because it doesn't exist. It was made up so that cardiologists had a controlled number of MIs. It has nothing to do with patients. And if you look at -- you want a number that predicts outcomes it is somewhere around the 60th percentile. So the number I want to know in my practice is how many patients am I going to send home who are harmed. So what is that number? There is going to be an acute outcome and a chronic outcome based on some number. And I'll sit down.

DR. APPLE: So, I'll answer that question specifically because there was a meeting the ESC and AACC had in east France and Allan Jaffe was there and I was the only laboratorian among cardiologists. And the world we live in, the laboratory world works in 95.5 percentiles. But the day that the troponin Assays were being starting to get FDA cleared back in the last 90s the imprecision was horrific. All right. So it was a CKMB assay in a troponin form. So the cardiologists rightly said we can't afford the false positives that we're going to see because of the noise around that

cut-off at the 97.5, so instead of using two standard deviations it was bumped to three. So I will propose that down the road as these assays are cleared and the precision gets improved likely there will be a discussion to move it back to the world we grew up in at the 97.5th like everyone else has it defined reference range. That was the history am I correct, Allan Jaffe?

DR. JAFFE: Absolutely.

DR. APPLE: Absolutely, that was the reason because that extra one and a half percent they didn't want to be categorizing patients as MI if they knew it was noise around the assay. That was the history of the 99th percentile. And it’s stuck since then.

DR. CHRISTENSON: Yeah, so I think that was such a great contribution back in 2000 to come up with something that made it a more sensitive assay. But I think what Dr. Peacock asked is really a very important point. And at that time good was -- I mean it was good to say the 99th percentile but it was perfect was the enemy of good at that time because there was no evidence that really showed this. Now and this is why

I am so interested to hear about the adjudication process there was no gold standard for and we're talking now about -- our question now has to do with myocardial infarction, not pulmonary embolism, not heart failure or others. So I think what we have to do is really hone it down to the questions and then as was said earlier by our statistical colleague in the intended population look at the cut points and I think maybe that is what Frank is driving at risk in that intended population, not necessarily in college students or all-comers or whatever the 99th percentile would be.

DR. GREENE: I also think it is a great point and I mean I've never heard an endocrinologist ask me what percentile the cut-off for glucose is or for hemoglobin A1C. And so again tying in the adjudication and the clinical outcomes is key and while the 99th percentile is really important because of the lack of standardization for troponin assays I think that getting to that question is the only way that we'll know what the proper cut-off should be.

DR. SAENGER: And I think just to reiterate

that point. I mean that's why there is a push to be able to have so many kind of detectable or quantifiable troponins at that low end because usually for every other lab assay that we have we can measure something. We see a gaussian distribution or maybe it is skewed but we can measure something and then we can derive a reference interval off that. For the contemporary assays of course we can't see anything below this whatever arbitrary 99th percentile there is. With high sensitivity assays it depends of course on the assay or platform and it is a little bit different. But that's why there is a push so we can more accurately derive what that reference interval should be.

DR. BECK: I'm going to go to the lady in the back.

MS. AJONGWEN: My name is Patience Ajongwen. Based on all the conversation that you have been having I want to go back to the sample size. I've heard the minimum of 300, I've heard larger sample sizes which uses confidence interval above variability. My question is based on what Paula stated the manufacturers have the option to do an overall 99th

percentile in that case the minimum of 300/300 gives you 600 total for a gender specific cut point. So I want to throw it back to the panel when we are talking about sample size here are we going in the direction of the gender specific? In that case do we know what the subgroup is enough or are we talking about the 99th percentile? And keep in mind depending on whether we are using the intended use population as our statisticians say it might narrow the 99th percentile. I just want to understand Dr. Christenson threw out the minimum what is the context in the sample size in the context of all overall versus gender specific?

DR. APPLE: I think from the IFCCs when they first came out we were talking about 300 men, 300 women after exclusions. That was for the 99th percentile as Rob pointed out. I think it is 385 if you want to get a 95 percent confidence interval. So we are talking specifically individual men 300 or whatever the number is and same equal number for women. And derive a sex specific cut-off for men and a sex specific cut-off for women. That is my view from our task force and from the AACC Academy is what we're recommending.

DR. BECK: Go ahead.

DR. CHRISTENSON: And again the problem with that approach is that you have very big confidence intervals. You have very large confidence intervals. So if you do it in one normal population over here in this area and then you do a separate study in another normal population using the same criteria your cut-offs are likely to change substantially because of that 90 percent confidence interval. That's the danger of using -- or that is the issue with using a relatively small number like 300 to define a cut point.

DR. GREENE: But I do think we are all in agreement that you do need to analyze them in a sex specific manner.

DR. PANTOJA-GALICIA: I have a quick question. Is there a clinical meaningful when you obtain a confidence interval are you looking for a special width of the confidence interval because that's when I determine --?

DR. CHRISTENSON: Yeah, I mean that would have to be probably some sort of consensus. I mean we'll know it when we see it; right? But we know that if it

doubles, if it is double the value, if it is ten and 20 as the confidence interval we know that that is too big. So how big should it be? We could maybe define a percent. Maybe that is something that a consensus group should come up with and that will drive the N. So it is a great point.

DR. BISHOP: I can't pretend to speak for all manufacturers but I'll speak for one manufacturer of a high sensitivity assay and kind of my perspective on a couple of things that Dr. Apple said and then Dr. Lias. You know Dr. Apple talked about it would be nice if all the manufacturers would get together and do something. I think for matters of business that's not really practical. It sounds like a good idea but it is not really practical.

And in terms of submission to the FDA any one individual manufacturer doesn't necessarily hear or see the feedback that the FDA has given to another. And so I think the FDA needs to be a little more prescriptive and a little more vocal on terms of what they are seeing and what they want to see. Otherwise each manufacturer is going to have to learn the lesson the

slow and the hard way.

DR. APPLE: Can I just answer that.

DR. BISHOP: Yes.

DR. APPLE: Just a quick answer that so I -- the concept I agree, don't disagree. But I think it would be important for the FDA because we do studies in my lab and I do four different studies right now in my lab and there are three different FDA individuals. And the messages I see to the companies are different maybe two out of the four. So I think that is a very important point that everyone gets on the same page.

DR. BISHOP: No, I agree. It is important that everyone is on the same page. But asking manufacturers to start talking with each other is not probably the way that is going to happen.

And then another comment I wanted to make about something that Dr. Lias said in terms of not being prescriptive about what population is enrolled. I think one of the challenges again speaking as one manufacturer of a high sensitivity assay is as assays get more sensitive the precision at the low end will get better and that will lower the cut-off. And as you

start being able to exclude subclinical disease that will lower the cut-off. All of these things are moving forward in such ways that are lowering the cut-offs and that if you define going back to something that Dr. Peacock said if you define the 99th percentile as myocardial infarction then the specificity of assays is only going to get worse and worse if you stick with that definition of MI as anything above the 99th percentile. And so I think manufacturers would be reluctant because ultimately there is some judgment on the assay whether it gets cleared or not and that is largely based on sensitivity and specificity. And so we're headed down a path where assays are getting better and the specificity is going to get worse if we continue to believe that anything above the 99th percentile is an MI. And so I think that is really the issue that needs to be resolved.

DR. BECK: Okay. We only have ten minutes left. So let's go ahead and the gentleman in the front had a comment.

DR. McCORD: Yeah, Jim McCord, Henry Ford, Detroit. And just to echo some of this overemphasis on

the 99th percentile I think sometimes we argue or discuss about this way too much. I have plenty of patients that are two, three, four, five times above the 99th percentile that don't have an MI. We're always looking for the next value in the delta, so the delta by far is much more important than the cut-off. There really is no "the cut-off" for an MI.

MS. BECK: Thank you. Courtney do you have a comment you want to make.

DR. LIAS: I just again appreciate the feedback. We are actually happy to consider any type of cut-offs. I think Paula said that before. So you know if manufacturers want to come with deltas they just need to show how well they work.

The comment on specificity also you know we are interested in providing assays that where the benefits outweigh the risks. And if assays are beneficial even though specificity gets worse, if that doesn't increase risk or provide new risk that can't be mitigated in some way either by labeling or by the way that they are used in hospitals, that might be acceptable. So manufacturers shouldn’t let that hold

them back.

With respect to working with companies on 99th percentiles we are happy to provide guidance on this question. What we are requesting is some interaction with the clinical community to help us understand what type of guidance to write. So we are happy, we are looking forward to seeing what your group or the IFCC group that you worked on has published. And we are interested in talking with people who have different perspective on that point. What we are not hearing yet is consensus. And so the more we understand where that consensus lies or if there is no consensus what the different points of view are the easier it will be for us to write that guidance.

We appreciate this discussion.

MS. BECK: Go ahead.

MR. JAFFE: Al Jaffe, let me endorse what Fred said about what we did with the 99th percentile at the time we had no idea where normal was. But let me also suggest that we've got to make sure that we're thinking of the same thing. And everybody has a little bit of a different need here. What we in the universal

definition are doing is defining when an MI has occurred. That is not defining risk. Risk is a totally different circumstance. In certain circumstances with high sensitivity it is going to be way down in the noise and that will be a valuable contribution eventually. And in some patients even having a higher value may not be associated with risk. It depends upon the clinical situation. So I think if we conflate prognosis and diagnosis we make this a much more complicated titration. And the other point to make that is important is as you start looking at subsets whether it is age, left ventricular hypertrophy, heart failure, the number of perturbations are gargantuan so staying with a normal population whatever way you wish to define it is actually easier because otherwise you'll have a different cut-off for Fred as he turns 60, for me as I'm a little bit older than he is, for anybody who has a little bit of renal dysfunction, for anybody who has a little bit of left ventricular hypertrophy and on and on and on. So there is a desire for everybody to say let's make it real simple, just do this and only this. And it fails often

because we fail to look at these different subsets. So we want to simplify and consistency amongst what companies do would be really important and helpful. But the idea that you can mandate each one of these parameters is not in the cards.

DR. APPLE: And I just add to Allan and to the FDA is after you hear this later is we in America are behind the times with troponin assays. And I think if we can work with the manufacturers closer because they've been -- these are out there in other countries and they are working and you can read the evidence based literature on diagnostics and outcomes. It's working. So hopefully from the end of today the two groups can get together and hopefully some of these assays if they are submitted properly can get cleared so when I become a patient in ten year I'm using a high sensitivity assay, I hope.

DR. BECK: Thank you. Real quickly because we only have a few minutes. I just want to move on a little bit. So one thing we are hearing is there is a lack of consensus. So to help address that we have people try to put this in the labeling so that it is at

least clear to the clinicians and laboratorians what's been done. So what kind of information should be in the labeling to be helpful to you? And Dina or Amy, I know you've --

DR. SAENGER: Yeah, I guess I would just say I mean for labs that are basically just going to be verifying the package insert reference range which they are required to do typically you would need about 20 minimum you know normal and you try to mimic whatever is in the package insert. So if the normal were excluded using NT-proBNP of 125 or you know EGFR of 60 or some -- it just helps the burden on the labs when they are trying to verify their own reference interval. So being specific about that is also very practical.

DR. GREENE: Also the outliers were touched on briefly by Dr. Christenson but I'll say that the sample size, the number of outliers that were detected, how those outliers were defined and then how they were either analytically or clinically further worked up is very important. So even if it was just this sample was this fold above the 99th percentile, we determined it was an outlier and when we tested it on an alternative

platform it was undetectable, we think this was an analytical error. Or on other platforms and then the sample also tested positive for cocaine and so it was excluded or something like that. But saying how many outliers were detected because that's then going to when we're either verifying a reference interval or we're getting calls from our clinicians we can better understand the limitations or why we might have these either analytically or clinically relevant false positive results.

DR. CHRISTENSON: And I would opine that transparency is the name of the game. So there are some journals where you have to put your data that you use to come up with your conclusions online. And then other investigators can take a look at that. So maybe that would be -- I mean that would be sort of the one end of the spectrum of transparency but that would allow folks to see what the outliers were that might have been thrown out. So it would allow the community I think to have a more informed discussion about what was done and what wasn't done. But being explicit about the population would also be important as my

colleagues have said.

DR. BECK: Thank you. That brings up -- oh, go ahead Norberto.

DR. PANTOJA-GALICIA: And I think it is important also to mention the method that was used to obtain the 99th percentile and the method that was used to obtain the confidence interval.

DR. BECK: Thank you. That does bring up we only have two minutes left but that does bring up the idea of outliers. So do you guys have any experience when you've had outliers in reference interval studies have you been able to identify clinical reasons for the results or when should we be allowing them to be excluded? Only if it is a known analytical error or when you can actually identify clinical reasons?

DR. APPLE: My simple answer is most of the time no. All right. So we do hundreds -- we go to health fairs in Minneapolis, St Paul area. We enroll patients. And you find outliers from the surrogate markers and first of all you lose contact with the people. And you are not supposed to feed back information because of IRB issues. But 99 out of 100

times you can't figure out why. It is a subclinical information is really nothing in a chart that says that they are diseased. So my experience is the answer to that is no.

DR. CHRISTENSON: The only thing I would say again is the patient an outlier or is the value an outlier. To determine if the patient is an outlier you need to do multiple measurements with other reliable assays to see if all of them, if you have consensus that all the values are high. If that is the only assay that is high and the others were low then it’s an outlier that is with that single assay. So I think that is the first thing you have to do and then you have to consult with a clinician to see if that value of the troponin is really something that should have been that it is pathological or not. So I don't know what the exact answer is but I think those two would be important points to determine.

DR. APPLE: Just that Rob and I are working with the universal sample bank. If you look at every assay that we have now eight assays for high sensitivity assays the outliers from one assay are not

the say outliers for the other. So I can't explain it.

DR. GREENE: But I mean often times even the samples that are defining even without the outliers, the samples that are defining the 99th percentile between assays are not the same if you take the same reference population. I don't know, I haven't seen that data but --

DR. BECK: Okay. Thank you. Real quickly just the two gentlemen in the back and then we'll end the panel.

DR. GORMAN: Thank you. I'm Bob Gorman. I'm with Siemens Healthcare. One thing I want us to not lose sight of is the shape of the distribution that's behind this. No matter what patient population you have as Jeff said there's a continuum there. These are extremely skewed distributions. Okay. So we also know that theoretically if you are at those extreme order statistics, those extreme percentiles they have larger variability than say the median. Okay. So you've got that working against you. Then you've got the actual shape of the distribution. If the distribution is so highly skewed then you are going to get wide confidence

intervals and that is actually just correctly reflecting the population and the shape of the distribution itself. So and your ability to correct for that is limited by the rate of the sample size calculation. Remember that the width of the confidence interval is always related to the square root of N not N. So we double our sample size it doesn't half the width of the confidence interval. So there is a law of diminishing returns that is going to be working against you. And again as these assays get more sensitive and we are able to look at whether it is the daily turnover of cells or you know somebody slips and falls on the floor and compresses their chest some more cells die. I think even if we had those healthy 20-year-olds you are going to see that. And so there's always going to be a skewed distribution. And so we are kind of stuck in the idea that statistics can only take you so far in terms of trying to limit the size of those confidence intervals. My point is that the confidence intervals are actually doing what they are supposed to. They are reflecting the variability in the patient-to-patient population.

MS. BECK: Thank you. Gene do you want to quick --

MR. PENNELLO: Well, I just have one comment. I'm Gene Pennello. I'm an FDA statistician. So in any given reference range study the distribution of the co-variants like age may vary from study to study and that could impact what the estimate of the 99th percentile is. One way to sort of calibrate all this across studies is to weight the data to some standard distribution that you are comfortable with. This is done I know from experience when they are estimating cancer rates across geographical regions they standardize to some standard distribution before they actually estimate the cancer rate. You can do the same thing here for any given reference range study. You standardize the estimate toward some standard distribution for age or other co-variant that you think is the right kind of normal population you are interested in.

DR. BECK: Thank you. And Dr. Jaffe and then we'll take a break.

DR. JAFFE: Let me just make the point that

sitting with a lot of these groups is as a clinician and with a partial laboratory medicine hat one of the tensions that always goes on is what is practical. If you want to get rid of outliers make your upfront criteria very rigorous. Many of us have noted and the literature would support it that if you wanted to add imaging, if you wanted to add the best imaging you could add MRI that you would end up reducing the number of outliers substantially. Nobody wants to do that. No one views that as a practical way to proceed. The same thing we fight all the time about how tight the restrictions ought to be on the various co-variants. How low do you have to be? I think in many instances they're too high. So we make a problem for ourselves by trying to find this balance between what is practical in the real world and what might be ideal for defining a normal population. I think that is the right thing to do. I think that is the real world. But then we can't sit up here and lament it and get unhappy about it. Those are compromises that we're all making. We need to acknowledge them and understand that they are going to cause uncomfortablenesses, outliers that

we're going to have to deal with.

DR. CAPOSINO: Thank you very much for participating in the panel. We're going to take a ten-minute break. So if we can be back here at 10:00 to start the second session that would be great.

Thank you very much.

BREAK

DR. CAPOSINO: Okay. I'm missing one of my panelist but I think I'm going to start.

During this session we will discuss clinical trials for troponin assays. We will discuss why clinical data is needed for troponin assays. Why more troponin devices are currently available outside of the United States. And we want to discuss how we can stimulate innovation of these devices in the U.S. What clinical trials for troponin assays typically look like? And some challenges that manufacturers have experienced with these trials.

We at the FDA are in the unique position to see multiple studies from multiple manufacturers and we cannot share the information that we review unless something is cleared. During this panel we would like

to share some observations from our review of these studies.

So why do we need clinical data for troponin assays? As mentioned in the previous discussion panel troponin assays are not standardized or harmonized. Tests results between assays are not interchangeable. And each assay has unique clinical cut-offs and different analytical performance. For example the precision at the cut-off can be very different from one assay to the next. Clinical studies are an effective way to estimate the clinical performance that clinicians and laboratorians can expect from each test and have always been provided to support troponin assays.

While all of our troponin assays are cleared as an aid in the diagnosis of MI some sponsors are also interested in different claims such as using troponin as a prognostic marker or to use troponin test results for the rapid rule out of MI. We do not object to these new clinical uses for troponin devices. Well-designed clinical studies to support these claims would provide reasonable estimates of the clinical

performance that clinicians and laboratorians would be able to expect for any of these additional claims.

Why are some assays not available in the U.S.? Since information supporting clinical validity is not needed for in-vitro diagnostics in some countries some assays that may not have been clinically validated might be available outside of the United States but are not yet available here. In the U.S. clinical validity information is needed to show that the device is valid, safe for use, and to provide reliable estimates of clinical performance so that laboratorians and clinicians have the information that they need to use these tests.

Another reason that some troponin assays may not be available in the United States is that a lot of sponsors have never submitted them to the FDA. And we are unable to clear devices that we have not reviewed.

Lastly in some cases the performance data that we have reviewed has shown some issues. Some assays that we've reviewed have demonstrated poor clinical performance. For example a test may have very poor clinical sensitivity and the number of patients with an

adjudicated diagnosis of MI that did not have a single test result with the investigational device above the clinical cut-off was more than 20 percent. For some tests high imprecision around the clinical cut-off can sometimes contribute to poor clinical performance. For some tests we've observed big clinical performance differences between recruiting sites when the testing is performed directly at the clinical sites that enroll patients. Some of these sites have had poor clinical performance that is not clinically meaningful. When there are many differences at the recruiting sites, for example, demographic differences in the recruited population, different types of samples that are collected at the sites, it is difficult to discern if the poor performance has something to do with the demographics, poor performance for that sample type, poor device performance, or some issue with the usability of the device. And this will be discussed in more detail during the point of care panel.

As these observations highlight, in some cases it is more obvious that there is a problem with the performance of the device for example when we observe

poor precision and poor clinical performance. However in other cases poor trial design and execution may also be contributing.

What do clinical trials to support the use of a troponin assay as an aid in the diagnosis of MI typically look like? Companies perform the cut-off determination study as discussed in the previous panel this can be the reference interval study if the sponsor is using the 99th percentile. Others also perform small pilot studies in the intended use population to derive cut-offs that are designed to meet clinical performance targets. These studies are very important. The better the cut-off determination study the more confidence a sponsor can have that they will successfully validate their test.

The validation trials enrolls an all-comers population. The term all-comers has led to some confusion. What we mean by an all-comers population is the subjects presenting to the emergency department where the treating clinician uses the troponin test in the overall assessment for MI. Serial samples are taken using the investigational device concurrently

with standard of care biomarker testing. The case reports from the subjects are submitted to central adjudication. And sponsors have used different approaches in their adjudication. Most sponsors use a group of three or five adjudicators reviewing all patient files and applying majority rule for final diagnosis. Some sponsors only submit patients for the third or fifth adjudicator when the first two or four do not agree. In these cases the tie breaker is not always the same adjudicator and is blinded to the fact that he or she is that tie breaker.

To address problems with throughput some sponsors have put together large pools of adjudicators. For example they will have 20 adjudicators in total and will create random panels of three to five members in order to adjudicate the files more quickly.

The ultimate goal is to show how well the test performs clinically using the adjudicated MI as clinical truth.

We've reviewed some studies from different manufacturers and today we would like to share some of the challenges with trial design that we have observed.

So first we want to acknowledge that these trials are difficult to design. For example we often hear that it can be difficult to obtain informed consent in a timely manner that would allow both for the timely enrollment of the patient and the timely collection of samples needed for these studies. We look forward to understanding how often this occurs and if there are strategies to improve the informed consent process for these trials.

We've observed trials where the clinical cut-off is not optimal and the clinical performance suffers as a result. We've observed potential bias that is introduced into the clinical study because of study design. As will be discussed in the next panel in some studies we've observed pre-analytical issues such as sample instability. For example for a hypothetical assay the samples are stable for two hours but the clinical protocol does not account for this and the samples are tested outside of their stability interval.

We've also observed trials where subjects are excluded from the trial and these patients are part of the intended use population of troponin assays. For

example a trial may exclude patients with previous hospitalization or previous MI. However, we understand that for troponin a treating physician will use the test to determine if any patient may be having an MI including patients with previous hospitalization or previous MI.

We've also observed issues with adjudication. For example the adjudication for a hypothetical assay may include very specific adjudication rules that may not be consistent with clinical guidelines or practice guidelines such as by including a very specific change in troponin concentration into the adjudication rules.

We also want to acknowledge that these trials are difficult to execute. Trial sites may not have the resources or training needed to perform these trials. While it is not unusual for trials to report some deviations and those are expected we've observed trials where so many deviations were reported that the results of the study were affected. A common deviation that we observed is that testing sites sometimes test samples outside of the claimed stability interval. Other common deviations result in missing data. For example

a hypothetical assay may include two different matrix types in the validation study but the study sites do not consistently collect both types of samples for each patient and as a result the trial may not contain sufficient clinical information from one or both sample types. This observation will be discussed in more detail during the point of care panel. Another issue that results in missing data is when testing sites do not collect investigational samples at the same time that they collect their standard of care draws. For example for a hypothetical assay the investigational samples were sometimes collected one to five hours and as late as 20 hours after the standard of care draws. In this case more than 50 percent of the baseline draws, and this is the draw taken soon after the patient presents at the emergency room, were missing using the investigational device. This may be more of an issue if a trial is designed to collect investigational samples at time points that are different from the standard sampling for troponin that is done routinely at that site.

These issues can often result in poor clinical

performance and are difficult to resolve.

We are looking forward to a fruitful discussion on strategies to avoid some of these pitfalls.

We've identified the following discussion topics for this panel: Discuss best practices for trial design to minimize challenges. Discuss best practices for adjudication. What information from a clinical trial do clinicians and laboratorians need to understand the clinical performance of a troponin device.

At this time I would like to open the discussion and ask the panelists to introduce themselves.

Courtney Lias will moderate this panel.

CLINICAL TRIAL DESIGN

DR. LIAS: Hi, I'm Courtney Lias. I'm the Director of the Division of Chemistry and Toxicology Devices here at FDA and I'm moderating this panel.

DR. deFILIPPI: My name is Chris deFilippi. I'm the Vice Chair of Academic Affairs for Inova Heart and Vascular Institute in Falls Church, Virginia. I am

a clinical cardiologist.

DR. JAFFE: I'm Allan Jaffe. I'm a clinical cardiologist with interest in acute myocardial infarction and also because of my research have a role in laboratory medicine at the Mayo Clinic.

DR. McCORD: I'm Jim McCord. I'm a non-invasive cardiologist from the Henry Ford Hospital in Detroit, Michigan and also the Cardiology Director of the Observation Units in the Henry Ford Health System.

DR. NOWAK: I'm an emergency physician. My name is Richard Nowak. I direct the clinical trials in the Department of Emergency Medicine at Henry Ford Hospital in Detroit and I'm a clinical professor of Emergency Medicine at Wayne State University in Detroit and at the University of Michigan in Ann Arbor and we have a particular interest in biomarkers in our ED.

DR. VUCETIC: I'm Zivjema Vucetic and I'm Medical Director for Ortho Clinical Diagnostics.

DR. CAPOSINO: Paula Caposino, I'm the Branch Chief for Cardio Renal Diagnostic Devices.

DR. WIENEKE: Hi, there, I'm Jacqueline Wieneke. I'm a holdover from the panel this morning so

I don't think I need to introduce myself. But I do look forward to a very vibrant discussion. I very much appreciated learning earlier about what the panel had to say. And I look forward to more at this time.

DR. LIAS: All right. Thank you. So I want to thank all the panelists for agreeing to discuss. Our goals for the panel are really as Paula stated. We tried to put together a diverse panel that might have either experience and/or interest in providing feedback to us and to others in the room about the challenges that people face in running these trials because we acknowledge that they are very difficult. And then hopefully at the end of the session we'll have some ideas about how to overcome some of those challenges to make these trials successful so that new and innovative, safe and effective troponin devices can reach patients.

So I'd like to start this off by sort of asking that general question. For those of you who either have been involved in troponin studies and/or studies like this that you can imagine what some of these difficulties might be what are some of the

pitfalls that you all see. We've described a few that we noticed in the submissions that have been sent to us. But what are some of the challenges that companies and PIs face when running trials for troponin. And if you have any questions related to that yourselves please include them.

DR. JAFFE: You should start.

DR. deFILIPPI: So I think what we are seeing with the introduction of the high sensitive troponin assays is a new rigor around doing a diagnostic study in the emergency department. You know and part of the sell to clinicians and ED physicians to embrace a high sensitive troponin is the ability to more rapidly detect and exclude a myocardial infarction when it is used for this diagnostic purpose. So that brings into play the ability then to include samples at a very early point in the patient's presentation. So do you want to try to tie that to as the literature has in several articles to the time of the onset of symptoms. Now that can be quite diffuse and difficult. Is it important to really time that to the time of presentation, not the time of the clinical blood draw

which can be quite variable in the emergency department? Or do we want to tie this to the standard of care blood draw which is from a pragmatic trialist standpoint that is the easiest thing to do. And with those issues comes the ability to recognize those patients very early in their course because of course patients who have ST segment elevation have already been triaged and generally are not included in these type of trials. How quickly can you recognize them and then how quickly can you get through the consent process and collect these samples.

DR. JAFFE: I'll make four points in regard to issues that I see. The first is that most of the trial or many of them and in particular the European ones have systematically lost a lot of the early patients. Take a look at TRAPID as one example, time from onset of chest pain to presentation was 1.9 hours but it took them an hour and a half to get the first sample. That is a 3.4 hour first sample and that's where you can start to analyze what the troponin is all about. So there is a terrible -- an important need to get the samples earlier in these individuals if one is going to

make a claim that they are going to be appropriate for early presenters.

Secondly, I've been concerned that in many trials the subjects have been honed down to those that are what are called patients whose primary concern is ruling in or ruling out myocardial infarction. And in some senses the way they do it in Europe where they pre-triage patients into such groups so that the incidents of myocardial infarction is 25 or 35% in some of those studies. That may make sense. In the United States where the frequency of MI in the ED is in the single digits most of the time. What it does is it develops a protocol for the evaluation of a very small subset of patients. And then the question is what do you do with the rest. So I very much think that we ought to include all-comers. If you don't include all-comers and you then do this sort of pre-screening the people who present atypically get left out. Who are they? They are the elderly and I'm getting there so I'm sensitive to that group. And women. And we then have a major controversy as to whether or not there are differences in the diagnostic performance of these

assays for women. So terribly important to get all-comers in in my opinion.

A third issue that is terribly important is that often these patients don't have late samples so that clinical practice because most of these are observational studies is to send these patients home if for whatever the reason because that's the local way of doing things; that's the protocol. Late samples become important if you look at even with high sensitivity there is an incidence of two, three, five percent that don't rule in until six hours. It makes sense, troponin release is blood flow dependent, you've got a total occlusion, you've got an infarction behind it, it is going to take time. So it is terribly important that late samples be gathered.

And finally in regard to adjudication I think adjudication can vary but it needs to be codified. You need to write down what it is and what the criteria will be whether or not it is going to be MI/no MI and sometimes you can't tell MI/no MI. Sometimes it ought to be MI/no MI, or we need some more clinical information from the hospitalization. But it has got

to be defined up front because if you don't define it upfront it can really become a slippery slope. And I suspect that's some of what the FDA sees.

So those are four things that I think are terribly important from the point of view of clinical trials and getting them right.

DR. McCORD: And with regards to clinical trials there's a lot of things we really could say, there are a lot of issues. I like to break this down into two big categories just to get an overall look at this situation. The future and what is really happening now. So I think as far as the future goes clinical trials with troponin are going to be quite different. We're going to be seeing these assays used more in the outpatient setting than the inpatient setting for general risk stratification for someone who might be in an adverse event who may benefit from a statin. There are trials out there to help predict stroke risk and atrial fibrillation and is being used commonly even now in regard to chemotherapeutic agents to help predict cardio toxicity and maybe intervene there. So how clinical trials are going to look in

that domain is going to be I think quite different and is still being thought about and studied and so forth. And I think in that situation we're going to be talking about very, very low levels and any change in that.

And then in regard to what I think our present topic mostly is in regard to someone who may be having an MI or not. Echo a little bit of the sentiment of Allan and Chris. We really need to enforce that these need to be all-comers trials and the definition of all-comers in my mind is anybody that the ED physician thinks may be having an MI, everybody who gets a troponin which is in some studies 15 to 20 percent of people coming into the ED and only excluding patients that have STEMI because in STEMI we really don't use the markers very much. And some of these studies sort of they have a lot of STEMI patients, really can give a misperception of how these would perform in the relevant population, the ones where an EKG is not diagnostic.

And then in regard to adjudication I think that is a critical piece. And to give a very little short anecdote after the second universal definition of

MI came out I read that and I was a little confused with the document. So I sent a case to two of the authors from that document and said hey would this case be a type-1 or a type-2 MI and one of those authors is in the audience who I will not actually name and there was disagreement even from the authors in regard to if this would be a type-1 or type-2 MI. So to echo Allan's sentiment exactly how you go through adjudication process I think there needs to be a simple maybe one-page document on how you apply the universal definition of MI because not to be too critical but there's some ambiguity in my mind in that document.

DR. NOWAK: I think as Jim said about, Frank would tell you I think 15 percent of all patients that come in to the ER get a troponin drawn on them. So if there are 130 million people seen in ERs, so maybe I don't know 30, 40 million people get troponins drawn. And I guess if you suggest that they get serial ones, I don't know maybe 40 or 50 million samples of troponin results are seen by ER physicians. So it is actually quite different for the ER physicians is that they see all this broad spectrum of the use of troponin in the

ED or the misuse or I don't know what the hell to do with it in the ED and sometimes a cardiologist can't be of any help. So when I look at the -- I see troponins all day long, all day long and try to figure out what to do with them.

From an FDA perspective the rule -- the aid in the rule in of AMI is very helpful. I think it would be very helpful if the FDA said that they were going to change that to verbiage that said troponins are going to be approved for the rapid rule out of AMI and the making of the diagnosis of AMI because that would actually give incentive for people to upfront help out in those that can be ruled out in a very early manner. And that would be exceptionally healthy for the emergency departments. We as opposed to Europe have a very broad net. If you think you had chest pain a year ago you'll get a troponin drawn. It is just very broad and I think part of it is the legal issue that in missing an AMI here in the U.S. is a problem. You get sued and if it is atypical they'll say well, it was atypical, don't you know you have atypical AMIs, why didn't you do something. So we do that.

So I think if the FDA actually encouraged newer trials to develop a strategy, prospective strategy as to whether or not they are going to use the LoD, they're going to use a 30 minute, a one hour rule out or whatever but actually prospectively evaluate that so that when the product comes on the market people can feel much more comfortable about discharging or at least looking at other diagnoses very early on in the ED evaluation.

The other thing I think and I think this is going to make it more and more challenging is I see troponins all day. They may not be going up. They may not be going down. They may be stable. So what does it mean in all these other diseases? And cardiologists generally don't get involved in those because that's an ER thing. So actually I think what would be good is if troponin trials were being designed that you would have all the patients adjudicated into different buckets. One would be ACS, one would be non-cardiac ACS disease, one would be non-cardiac disease and one might be we don’t know what you had. And I say that in the sense that because if you started looking at all of those

within those groups there will be some that have troponins that are elevated and some that don't. We kind of know that if you have an elevated troponin whatever you have it is not good long-term. But I think it might start to shed some light on what to do with a patient in the ED who has a troponin with a similar disease of someone that doesn't. For example if someone comes in and has a hypertensive urgency and has a non-detectable troponin and another one has a detectable troponin maybe even at a low level what's the difference? What should the treatment be? Short term, long term? And so I think you have an opportunity when you put all these people into a troponin trial to glean out a ton more information than you have now with just trying to rule in AMI from an FDA perspective. I'd like to see them more formally pursue a rule-out prognostic, a rule-out prospective part of the trial. And then to start to get into some of these other issues because you have the patients, you have the troponins, you have the other diseases. And I think as we sort that out I think it would result in much better patient care. And it would be easier for the ER people

to figure out what to do with these numbers when they see them.

Thank you.

DR. VUCETIC: So I'll be talking a little bit about the manufacturers' perspective at least from one manufacturer's perspective. And I really appreciate all that you've said and that's extremely relevant to what we are doing. I think for us, we're kind of in between both the clinical practice and the regulatory aspect of clinical trials and designing and running them. So when we think about designing the trial first what we look is what is going to be clinically relevant. So we want to incorporate everything that you said, you know from looking at the relevant population which is an ED population that is really hard to not only enroll but to consent. We have multiple time points that we need to collect for these types of trials. And we also have the adjudication part where looking at what is the gold standard diagnosis and how do we normalize across all of the sites that are enrolling into the study. These are I think the three most important things for us.

But on the other side I think because we are working within the framework of the 510K process we always look what is our predicate device and how does this influence how it's our device that we're bringing out with this new clinical trial approach is going to be assessed. So it is kind of difficult for us to really be innovative and going forward if we have some limitations or if we don't really understand how are these novel things that we are including in the clinical trial going to be evaluated and assessed and how's the substantial equivalence is going to be determined. So I think these are kind of -- it puts a little bit of a boundary on the manufacturer in the design of the clinical trial.

I think --

DR. LIAS: Could you speak a little closer to the mike, please?

DR. VUCETIC: Yes. I think one of the examples for that is really for us collection of multiple time points and how early do we want to go. I think there are currently no, at least as far as I know, there are no devices on the market that have really early time

points. So if we are trying to design a study that collects samples at presentation and then you know an hour later or three hours later what is our best strategy to do that? Is it really collecting time points as the standard of care? Or should we really set the times later on? And how is this going to really be evaluated based on what's already on the market?

So I think there is -- I'm really looking forward to both the opinion from the clinical practice and the FDA in terms of that.

DR. CAPOSINO: So from our perspective is we are very interested in promoting the use of the trials -- sorry, how about now. So we are very interested in promoting innovative troponin devices that the clinical community needs. And I think we understand that it’s difficult to be the first to come and do that. And I think in this way figuring out what those trials should be because they may be a little different if you are trying to rule in as opposed to rule out the trial may be a little bit different and to take that into consideration so that you are able to design and

execute the trial that's going to get you your successful device. And I think that is what we want to talk about today is how do we get there? How do we make these easier to do? A lot of these companies are not seasoned clinical trial people. And we just want to acknowledge that the challenges that they see may in part be by the way these trials were designed.

DR. WIENEKE: Yeah, everyone has made some really great points and from my perspective from the clinical side the FDA raised a lot of questions as far as some of the submissions that do come through.

I do want to just make a couple of points that I don't see my role at least at the FDA as telling the clinicians as to how they should use their devices or what clinical trials would support how they want to use it. So we are definitely open to the clinicians with their input. And if an intended use as a rule out is of more benefit than a rule in I mean technically what we -- I believe what the intended use is, is it's an aid in the diagnosis. So it doesn't say rule in or rule out, it is an aid in the diagnosis. If there is a better intended use that would be more useful to

clinicians, I think we are open to that. I think what we have today is not a perfect trial design with this all-comers trial design. It is the best we have right now. But we certainly are open to whatever clinical trial design is of most benefit to the clinicians. That's our role is to get the best devices out to the clinicians so they can take care of the patients the best they can.

So again aid in the diagnosis is where we are. If there is a better intended use please let us know.

The all-comers part is one of the most important things and again I guess right now speaking to the manufacturers if you use specific exclusion criteria to enroll your patients then those patients are technically not part of your intended use population. So what we are trying to get with the all-comers study is what patient population is the test being used in. From our perspective it is a pretty broad population, like Dr. Nowak said. So from our perspective that's what we need to see. If you are excluding patients based on this criteria or that, just be aware that that is no longer part of your intended

use population and we're going to have to kind of put it somewhere and tell the clinicians that the test was not evaluated in that patient population.

And then the last thing I would like to actually discuss because we have had a couple of submissions. Regarding this whether you are including or excluding STEMI patients. So and again it just might be that I'm naïve as to how it is actually used, how troponin is being used these days, in my opinion or at least in my experience troponin is being ordered in STEMI patients. If it is not please let us know and we have had I think a couple of submissions that exclude those patients but that's sort of an issue for me if troponin is actually being used in those patients. I need to see how the device is used in those patients if it actually is being ordered.

DR. LIAS: I think that is a great question. Let's come back to that. I know Allan had a comment and I've written down the STEMI question for us to come back to and I have a few questions also from what we've just heard. And we'll get to questions from the audience in a little while.

DR. JAFFE: Yeah, I just wanted to make a comment about my two colleagues to my left. I think we have to think through what it is we really want as the FDA role. And Frank wants to know who's at risk. You really want the FDA to make the clinical guidelines to define what's risk for you? I would argue that we probably don't. I'm not sure, as enthusiastic as I am about long-term studies looking at how we use troponin at low levels and how interesting they are, an interest of mine. But I'm not sure that we want the FDA to be the eventual determinater of what protocols we use, what the cut-offs are, what the deltas are. I think that is asking for them to do what we who make clinical guidelines ought to be doing. So I think what we want the FDA to do is to look at assays and say do they work, and the way in which they're said to work, and are they safe. And the rest of these details as nice as they might be and I understand Richard's concern and Jim's concern; we've talked about this before, are things we ought to do and take over from the point of view of the American College of Cardiology, the universal definitions.

DR. LIAS: I appreciate that perspective and we would love that help. Our panel today is to help us understand how to as a community FDA, manufacturers, academic PIs, laboratories and physicians run these trials in a way that is practical so that we can get those estimates, so we can get that information to put in a label to say how well does this test work, how can you expect this to work in your patients.

So the comments earlier were really helpful and I think all boil down to a few of the things that you all summarized. And so I want to delve a little bit into some of the specific questions. One of the things that you all seemed to agree on was the necessity of using this all-comers population, the population of people who would get a troponin test ordered on them in the emergency department. So how can we better design these trials or run the trials or execute the trials to improve the enrollment of that intended use population?

DR. deFILIPPI: I can try to answer that. So I think you can have broad inclusion criteria, for example, all levels of renal disease, trauma and so forth. But I think behind the scenes you have to look

at the sites that are being selected to make sure that they represent an appropriate demographic, that is represents the broad demographic of the United States. Inner City, academic type hospitals, community hospitals where you often see higher rates of inclusion of myocardial infarction. So that's probably more subjective but also is critically important when you design an all-comers population.

DR. LIAS: So what do you think are the reasons we can't get them now. So you know the trials right now are designed typically fairly broadly to include this population. That's what is indicated in the trials that we see. But I'm hearing some feedback that that sometimes those are not the patients consented, those are not the patients that are actually included when the trial is run. So what could be done given that a trial has a broad inclusion criteria, not many exclusions, what can be done to actually get and I think I heard there were a lot of early patients lost. I think that is what Dr. Jaffe said. So how can we improve the enrollment of the patients that we really need to see in these trials to for example improve the

ability to assess these markers for rule out.

DR. JAFFE: One of the reasons the early patients are often missed is because of the facility of getting the research sample. Now various places are different. The IRB at the Mayo Clinic allows us to get verbal consent to get that first sample and then come back. Other places may or may not. So I think it is incumbent upon people doing the studies from the point of view of getting early patients in to ask their sites whether or not they have the ability to get samples early after the patient presents. And that should be part of the criteria of getting into such a trial so there are adequate numbers of these patients. The vast majority of non-STEMI patients don't present early. They mostly present late and that's why all of these magic rapid rule-outs work so well because you're already three to five hours down the road. But the early patients then that they are going to be applied to become terribly important. So one way would be to screen your sites upfront and to suggest and embrace some of the more liberal IRB policies which some places are allowing.

In regard to critically ill patients they often are not included because they are not thought to be part of the intended use rightly or wrongly. So a patient who is septic where some EDs they may get a troponin. If they do I think you've got to include those people. If they don't then you don't because I think otherwise you are picking and choosing. So I think an all-comers population is simply said you get a troponin and you're in.

DR. LIAS: Yeah.

DR. JAFFE: In regard to STEMIs --

DR. LIAS: Okay.

DR. JAFFE: -- just quickly. Two things happen in regard to STEMIs. First of all many STEMIs even bypass the ED. They come in through EMS, identified as having by an outside electrocardiogram and go right to the cath lab. That is a common practice in many places. In addition there is a very rapid triage of those patients in the emergency room that makes it hard for a research enterprise to interdigitate with those people.

And finally depending upon the expertise

locally there may or may not be a lot of yield for using a troponin. The diagnosis should be mostly clinical. We certainly use troponins. We certainly get troponins in those patients. But usually they are not indigenous to the decision that leads to go forward for cath.

DR. LIAS: So I think one of the questions though should those people be excluded. So maybe -- it sounds like maybe they may not be included easily but should they be excluded?

DR. JAFFE: If they're sent emergently for a procedure based on their electrocardiogram and clinical presentation as a STEMI they should be excluded. Now you have to be careful because there are some sick non-STEMIs that go rapidly to the cath lab that I think you do want to capture.

DR. McCORD: And of course all patients who come to the emergency room who may have ST elevation where the suspicion is indeterminate and so you would want to include those because that is where the troponin could aid in the diagnosis.

DR. WIENEKE: And so my question as far as the

clinical trial design should the STEMI be specifically excluded from the criteria. In other words should the protocol indicate exclusion criteria all patients who present with STEMI?

DR. JAFFE: I think you could do it either way. You could include them as a post hoc defined subset if you want it. They're very logistically difficult to deal with. They are not the most common circumstance in which clinicians feel the need for troponin evaluation because it is done on so many other clinical criteria. So from my point of view leaving them out is not a major problem. On the other hand I wouldn't object to leaving them in but I'd put them in a separate category. Otherwise their troponin values which generally are quite a good deal higher and particularly after reperfusion are going to skew the metrics that you end up with.

DR. LIAS: Another option I think I just heard was some of them are easily just not -- troponin isn't needed but that sometimes if it is needed they could be included; is that what I heard.

DR. McCORD: From a practical perspective if it

is a clear STEMI they are kind of excluded in other words, the blood, they are running off, they are getting blood samples but no one is actually checking them, they are getting the cath done. It is clearly very helpful if it is a STEMI in retrospect as we say it sometimes where the marker comes back and now we look at the EKG again, they go oh, actually this guy should have gone to the cath lab. So I think a lot of them get excluded kind of anyway from early sampling. But I agree with Allan's idea that you can have them but they are a separate group. The key question is how the markers really play out in regards to patients having a non-diagnostic EKG.

DR. LIAS: In the interest of time I want to get a couple of these questions so we can get to the audience comments. So we've heard a little bit about the desire for rule out. As you all know there are no perfect tests. So when you set a cut-off you often have to trade off sensitivity with specificity. Rule out tests often have to have a high negative predictive value. If you ever want to use it in that way. So either you would have to have a test potentially with

two different cut-offs or you would have to have a test specifically designed for rule out or in the literature we've seen the options of adding another marker to assist in rule out along with troponin.

So what needs to happen next for getting to the situations where the emergency department physicians get tests that they would like to have for rule out? Do they need to have new tests with new cut-offs? Are you willing to trade off some of the other performance parameters?

DR. NOWAK: So actually there are reports in the literature that already give you pretty good direction. So the assays are getting so sensitive that it really does appear that if you are below the level of detection when you present it is hard to understand how your symptoms could be related to an acute MI when we don't see a molecular level bump in your troponin. I mean it just doesn't make physiologic sense. So I think you -- the LoD I think is going to be very big and then very early changes. If you have very sensitive troponin you have in MI you'll see some change very early on. So if in an hour you don't see

any change, if it is low and it remains the same you could probably I think say that is not an AMI. And people are using those but I think it would be important for manufacturers to actually incorporate those things in their studies and then actually to probably get an approval for that which I think is stronger than just --

DR. LIAS: And then yeah for a rule out claim it comes to the second point is how do you get those early patients? So --

DR. NOWAK: Well, the early patients I mean you can -- if you decide whatever --

DR. LIAS: We already covered that topic.

DR. NOWAK: Whatever Allan decides is early so if Allan said well --

DR. JAFFE: Two hours.

DR. NOWAK: -- if you've come in within an hour or an hour and a half and we think it is not reliable that is an easy thing for ER guys to understand. You know this rule out works unless you've had symptoms for less than an hour. Having said that I'll say when you start talking to people about their

symptoms and how accurate they are and the concept changes and whether it comes and goes, it is not so simple. But I think if you have any concern that it is too early what do you say, not applicable to this patient, wait another hour and do it and you are done.

DR. JAFFE: I think that is the key which is I think the reason the values that are very low work is because and particularly if you subset patients as low risk with a low risk ECG as been shown by Yader and Fred and many others and you have a low troponin since almost all of the co-morbidities that lead to cardiac disease are associated with rises, albeit within the normal range. A very low value says a third low risk, that is a really low risk patient. I would be reluctant to do that in a high-risk patient even if they had paradoxically a low or a very early patient. So I think the LoD is important. I'm worried about the one-hour rule out because biomarker release is blood flow dependent and there are going to be some where you won't see that and the tolerances for example in the EFC guidelines of the difference between three and five are beyond the ability of the assay in most

circumstances to achieve it. So I think we've come to wanting everything a little bit too fast and we need to be very conservative to make sure that when we apply this to patients we don't hurt a large segment of patients because there may be ten patients in that very early group and they cut it at four because that is what it looked the clinical data suggested but that's not proof that it is safe in that group in my opinion.

DR. deFILIPPI: I want to add one comment to that so if you really want a true early rule out at a very low level and I know some manufacturers have kicked this idea around it is a different type of trial. I mean most of these trials we talk about are prospective blood collections and there's no impact on clinical care. You would want to do a prospective interventional type of trial in which you take that result and act on it or you don't act on it. Once you've got there's a lot of data in the literature to support that a patient would do very well. And then you could make the claim at this very low level these patients had a very good outcome. And probably otherwise they're still going through all the standard

care, getting multiple blood draws, getting a prolonged ED stay, so forth. But I think you need to think about designing interventional type trial design.

DR. LIAS: Thank you. I'd like to go out to the audience now. We have a few people who have comments.

DR. APPLE: Fred Apple. I was going to address the point of trial design. Two, one comment and one suggestion.

So the designs that we do for companies they are a disaster and I'll tell you why. Currently because you have to have 24/7 coverage, you have to have dedicated nurses or someone following the patients. It is so cost prohibitive. And you miss patients. If someone takes a bathroom break and you can't get that time specimen. So I think that concept has been very, very difficult. And it is tough to enroll as we've talked about, you've all talked about those consecutive patients. It is never consecutive like the all-comers.

So a proposal I'll throw out is two things what we used to do in the old days is to bank every specimen that came in and you freeze them. It is

designed prospectively but it is a retrospective analysis. So we could actually take the data of everyone that had a troponin ordered by indication, put them in a biobank, the IRB, I checked with my IRB you can get an expedited or full review to use waste specimens and waste records. You bank those and then you go back and you get every troponin ordered.

The second way to do it is the way we do our investigator initiated studies. Let's say we did our UTROPIA study when we worked with, it was an investigator study with Abbott, is when the contemporary assay came off the instrument it reflects as a waste specimen right to the high sensitivity assay. We have fresh measurement samples. So if you design your study, if you are Siemens and you just take Siemens' sites that have your contemporary assay and as the result comes off the instrument it does your high sensitivity, you could probably take three or four sites and enroll 2,000 patients in two months and adjudicate along the time. The cost will be a huge savings. You won't have to get informed consent because they are waste and the only time you need

informed consent is if you need outcomes. And as Allan said most IRBs will allow you to walk up to the patient admitted and consent them for the IRB for an outcome after the samples have been banked. So I think it will allow other companies that have small budgets that can't afford to do these multimillion dollar, three to five million-dollar studies, that they have to pay all these sites to hire all these people to collect the samples but you can actually bank the as you go, measure the samples, and then do it. So I think that is something that I wish the FDA would consider as a real way to do these trial designs.

The last comment is if you work in a lab and you do these studies to get the time specimens it's really tough because most hospitals don't even have serial orders in place. One doctor might order it at one hour, one might do it two hours, one might do six hours. But to set up a uniform, set up if you don't have a serial time, we have serial times at 0, 2, 4, 6. It fits my research needs too. But if you are going to a place that does 0, 6, 12 to plug in a one and a two, you are missing a lot of samples. So you really have

to get organized to do that. And that's very costly.

So those are two ideas I'll throw at the FDA to consider avoiding the informed consent for the expedited way to get all-comers and then every indication gets a measurement.

DR. LIAS: Thank you.

DR. PEACOCK: Thank you. Frank Peacock. So I have big question but we hit the STEMI issue and I have to address it because I don't think the panel was clear on this. You cannot enroll STEMIs period, none; don't do it. And here's why. If it is a true STEMI it should be in the cath lab. If you are consenting them to be a part of a study it is unethical. My IRB will not approve it. You can't do it. Now if it's -- as McCord said if it is one of these I'm not sure it's a STEMI, it turns out to be a STEMI, that's fine. But to intentionally delay somebody getting a cath is wrong. The mortality rate for delayed cath is one percent every three minutes. How long does it take to sign a consent form? About 20 minutes. I just increased their mortality by seven percent. Not appropriate. And there is no use for a troponin in a STEMI. So

exclude them if you know it's a STEMI. If you don't know it's a STEMI, they turn out to be, it is fine to have them, but that's it. Okay.

As an aid to diagnosis if you are having a diagnosis it is a rule in. It is not a rule out. And what we really need is the rule out. And I know you guys see a ton of data so I want to tell you what's really happening out there with no data. So I got asked to provide expert testimony in a lawsuit and I looked at it and here's what goes on. A lady comes in, she's got chest pain, the doc who is a normal emergency doc, who's read a bunch of literature himself, says, oh, well, I just did a troponin. It is negative, she can go home. One and done. She is sent home, completes her MI, comes back two days later in florid heart failure. And you ask this doc, it is like were you stupid you sent her home with one troponin. He goes no, I read this literature that says right here that you can do a one troponin and it's low and they don't have it. I go that's Europe. They have better assays than we do. And the reality right now is the United States is about ten years behind the rest of the

globe in our troponin assays. We are the third world of troponin. We are Nigeria I can get a better assay in Nigeria.

And so this is the challenge every company that makes an assay has a better one than the one I get to use in Baylor at my hospital. And so the challenge is what we've created a situation where the perfect is the enemy of good. We're trying to make these assays so perfect they can't be coming in our country. So we don't got 'em. So I'm using assays that aren't that great and don't stand up to the world's literature, and the docs read the world's literature and apply them here because they think they have a high sensitivity assay.

I go to conferences all the time. I say how many of you have got a high sensitivity assay and two-thirds of the audience raise their hand. They don't have a high -- they think they do because somebody told them they did. So the challenge is why cannot we have a standard if your assay is better than the one you have selling right now, that's good enough because it’s an improvement. That is how car companies work. Every

year the car gets a little better. I'm using a ten-year-old assay. What did your car look like ten years ago? There was no airbag, it didn't have antilock brakes, and every year they got a little better. If we used a strategy of better than what we got we would be better than we are today.

DR. LIAS: Thank you for that perspective. I think we agree and we would like to use today to help everybody get on the same page about how to get there.

DR. SANDOVAL: Yader Sandoval, Mayo Clinic. So I have a couple of comments. I have a few thoughts on how we've done this within our research with Fred at the Cardio Biomarkers Trials Laboratory. The sort of things I see when I read the literature, when we reviewed peered reviewed publications and others. There's two issues that I would like to hear the panel’s thoughts.

So one is and I think this was briefly touched upon is the adjudication process. So the literature has a lot of heterogeneity right now in the endpoint for the diagnosis. Some studies the primary entries were type-1 myocardial infarction, all others are for

both types of myocardial infarction, type-1 and type-2. Others actually look for acute coronary syndrome and if you look at other studies for example, Adapt, they use major adverse cardiac events. Should there be a consensus and what should that -- what is it we are looking for to stratify ACS, MI, type-1, type-2. The literature there is wide variability in that.

DR. LIAS: So before you get to the second point, why don't you want to hear the panel discuss that?

DR. SANDOVAL: What?

DR. JAFFE: I'm not sure that I know how to distinguish type-1 from type-2 MI all the time. And I write these guidelines. So the problem is not in the theory or the pathogenesis but the operationalization. So the first thing to say is I am as a reviewer when I review manuscripts terribly cynical about people who take the idea that they can absolutely detect type-1 versus type-2 MI. So I would say that you have to include both as part of your adjudication. I would suggest that the FDA not get immersed in trying to make the distinction of one versus the other.

On the other hand I do think there is room for variability when it comes to adjudication if it is written down. And I think what needs to happen is a plan needs to be written down and you could say, we're going to define ACS because we're interested long-term in whether or not there is some unstable angina in the patients and what their prognosis is and whether or not we missed them. Fine with me. Or you could say no, we're just interested in the diagnosis of AMI and we want AMI/no AMI. That's fine with me. I think there can be variability. I think the problem is that so often people don't write it down, then it becomes a slippery slope.

So I'm for flexibility but I would say an MI has to be called an MI and you shouldn't try and subset them at least not from the regulatory perspective. Now we could argue whether or not guidelines might do that but that is for a different group.

DR. LIAS: Thank you. So we only have about two minutes left and we have two more questions.

DR. SANDOVAL: Let me add two other aspects. So that was for the diagnosis, right. And just a

second aspect that it applies to clinical trials where we are assessing whether the strategies are good enough for implementation. There is also a lot of heterogeneity in the endpoint for safety. So some are just the index diagnosis during that index presentation otherwise our 30-day MI or the ones for 30-day MACE, these are two aspects, the adjudication of events and what is the window of safety that we're trying to address. A lot of heterogeneity in how the studies are being performed.

DR. JAFFE: One of the things I've been terribly critical about is the literature has failed to address what happens to patients who rule out. So if you take a look at the New Zealand/Australia example they have very, very good primary care. And if you go back and you force as an editor as I did some years ago them to articulate what goes on over 80% of those patients who were ruled out got investigated. I think it is probably very different from sites in the inner city where there is very poor care. So I think there is a legitimate role to look at 30-day outcomes but I think we have to do it in a better way by including

what is follow-up care as part of that analysis because that may explain a good deal of either the benefit or the detriment.

DR. LIAS: Thank you. Go ahead we have time for one question.

DR. deFILIPPI: I was just going to add that in addition to saying MI or not MI I would put a category of myocardial injury or not so the adjudicator is identified, the troponin is elevated they don't consider it a myocardial infarction so when you think of that intermediate or longer-term prognosis then that can be part of the follow up.

DR. LIAS: Great. We have time for only one more question. Any other questions please follow up with the panelists afterwards.

DR. GEE: Hi, I 'm Matt Gee from Siemens Healthineers. I just have a quick question because I heard there is some interest in or the importance of the LoD. And I just wanted to ask knowing that FDA's preference right now is the analytical measuring range to be based off the LoQ. I'm curious to know what this panel feels --

DR. LIAS: So we are going to cover this topic in the next panel. So I'm going to hold your question until then. And if you'd like to hear these panelists' perspective then please follow up with them afterwards.

That way we can get one question. I don't want to cross panels on different topics. Thank you.

DR. HUGHES: It's actually -- Karin Hughes, Astute Medical, it is actually not a question, it is more of a comment. Informed consent, so I think getting a verbal informed consent is not the norm. Just the process of consent alone for the patient or the patient's family to really understand what they are consenting to is a process. And in the ED it is difficult because you are in the emergency department, people are there for an acute issue. So that is an issue. It is also a matter of looking at inclusion exclusion criteria. So even if you have the best study coordinator in the world consent is not a three second process. So you asked why we missed early samples in our trials, that is a large portion of it. I think there is also confusion whether there should be or not from IRBs and what is allowed by FDA in terms of how to

get informed consent. So particularly if the patient can't consent themselves and often they can't you need to get a family member and unfortunately a lot of people who present to the ED don't always have a family member with them. So can you text, can you phone, can you -- and lots of times the IRBs don't know the answer to that question themselves. So we're trying to educate and ship them FDA guidance documents when they are available. So that all adds to the challenges of getting those really early samples.

With respect to leftover samples, thank God the common rule allows for that still but there are limitations on the manufacturer side as well and I think with FDA too. You guys are concerned if you have a new marker for the same type of indications you are worried about matrix effects, you're worried about stability. So banking samples makes it very difficult to answer those questions. And you also lots of time whether you like it or not you need data that is in the EMR, there is no way getting around that. You guys are going to ask the questions when we go to publish, what was this value, what was that value? Well, we don't

know because they were leftover specimens. And by virtue of that it is very difficult to get that information. And there is no way you are going to find that information out later.

So just I know it is not any answers but it is a lot from the manufacturers' perspective of how difficult it is in trying to get not just the right clinical people but to deal with all of those issues. And maybe some education of the IRBs may help as well so that the manufacturers aren't left to do that job too.

DR. LIAS: Thank you very much for your comments. We are out of time for our panel at the moment.

This has been an extraordinarily helpful panel. I know that personally I have heard some things that really help us understand some questions that we have and hopefully this can open up the lines of communication for further discussion on this.

Thank you very much.

Now we are going to head into our next panel.

DR. WELSH: Okay. It looks like our panel has

arrived. So I'll go ahead and get started with the next session.

Good morning everyone. Thank you for taking the time to participate in our public workshop on cardiac troponin assays.

My name is Kerry Welsh and I'm a reviewer in the Division of Chemistry and Toxicology Devices.

The purpose of this session is to discuss the Pre-Analytical and Analytical Considerations for Clinical Studies of troponin assays.

I will highlight some of our challenges with analytical issues we have seen in several submissions. Hypothetical data representing these challenges will be shown. I will also read the question we have for the panel on each topic.

This slide provides an overview of the topics to be discussed. Specifically the session will cover specimen stability, potential differences in specimen matrix, detection limits, analytical issues related to trial considerations and the emerging issue of biotin interference.

The first topic I will discuss is that of

specimen stability. Sponsors typically provide stability studies for two purposes. The first is to support routine specimen handling and processing conditions in the clinical laboratory. And the second reason is to support specimen handling conditions of banked clinical specimens used in clinical studies. Laboratories will want to know how long a specimen is stable at room temperature, under refrigerated conditions, the number of freeze-thaw cycles, and how long a specimen can be stored under frozen conditions. Specimens used in clinical studies may not be analyzed immediately and thus may be subjected to a variety of storage conditions. It is important to understand if the clinical performance of stored specimens will reflect what an intended user will see which is usually fresh specimens for troponin assays.

We have observed different troponin stability results depending on the assay used. As an example we have observed certain troponin assays that show instability at room temperature by two hours. This slide shows a hypothetical assay with a male specific cut-off of 25 nanogram per liter. As shown in the

first line of data in this table the baseline value is above the male cut-off but would give a negative result after two hours at room temperature. Likewise the other data in this table highlights that different values obtained after storage at room temperature for two hours may result in a different interpretation depending on where the assay cut-off is. While not shown on the slide this hypothetical assay showed nearly identical performance after storage at refrigerated and frozen conditions.

Thus a clinical study using this hypothetical troponin assay that is not stable at room temperature and did not account for this in the sample handling and processing conditions of the clinical study may result in biased clinical performance especially if samples are stored at room temperature for varying amounts of time.

Likewise some troponin assays have demonstrated stability for up to eight hours at room temperature but are not stable after freezing. This data table shows another hypothetical troponin assay that is not stable after freezing. If a cut-off of 17

nanogram per liter is used the data in line one would show a positive result that would be interpreted as negative if the study used frozen specimens. Likewise line two shows example data that may have a different interpretation than at baseline after a freeze-thaw cycle if a cut-off of 14 nanograms per liter as an example was used.

In summary we have seen different stability study results depending on the troponin assay studied. If a clinical study for troponin assay processed samples in a manner that stability studies suggest would be associated with specimen degradation this would raise questions about the validity of the clinical performance estimates.

The next topic I will discuss is the potential difference in troponin results that can sometimes occur in different tube types. Sponsors sometimes want to claim multiple matrices for their assay systems. Sponsors often choose to perform the clinical and analytical validation studies in one tube type such as lithium heparin tubes and use a matrix comparison study to transfer the performance claims to other tube types

such as serum.

The following table shows the regression analysis on a matrix comparison study for a sponsor who performed their clinical and analytical validation studies using lithium heparin samples and used a matrix comparison study to transfer the claims to serum specimens. The troponin results were nearly identical in the two tube types both by review of the line data and regression analysis that demonstrate a slope near one and an intercept near zero. Importantly the sponsor demonstrated clinical concordance between the two tube types at the assay cut-off.

However we have observed that some troponin assays show different performance in different matrices. This table shows some example results of the same troponin assay collected in different tube types. Line one for example shows a very different lithium heparin result from the serum with clot activator tube. If for example this assay had a cut-off of 19 nanogram per liter one may have a different clinical interpretation depending on the tube type. Likewise line two shows troponin results ranging from 19 to 28

which could result in a different interpretation depending on the tube type and where the assay cut-off is. While the example data in lines three and four would likely be considered elevated with most assay cut-offs these data show that different results may be obtained with different tube types which may be an issue for an intended user who is trending troponin results over time.

In summary the ability to use different matrices depends on the specific assay. If a troponin assay shows different results across matrices this may raise questions about the analytical and clinical performance of the assay based on the matrix.

We also sometimes see unusual results or flyers with certain tube types. This hypothetical data table shows examples of two specimens that are labeled as number one and two that were assayed with three different tube types, EDTA, serum and lithium heparin. Specimen one showed troponin results ranging from approximately eight to twelve nanograms per liter with EDTA or serum tube but had markedly elevated results in the lithium heparin tube. Specimen two show similar

findings. Observationally we have noticed we tend to see this phenomenon most often with lithium heparin tubes. We are not aware of any specific literature linking these flyers to lithium heparin tubes. We've seen some literature that has suggested these are due to micro clots or are perhaps analyzer issues. And we are interested in our panels experience with unusual troponin results.

The next topic I would like to discuss is that of detection limits. We understand that performance is important at the low end of the assay. We would like to highlight some analytical challenges we have observed regarding troponin assay detection limits.

First, I will discuss the typical experiments performed by sponsors to determine detection limits of troponin assays to highlight some of these analytical challenges regarding detection limits. I will first discuss the concept of the limit of blank. Sponsors typically define the limit of blank as the concentration that is only exceeded five percent of the time by a blank measurement. This is usually established by measuring a sample with no analyte 60

times with at least two reagent lots.

The graph depicts the limit below which 95% of all blank test results fall. The purple area of the graph shows the probability that the blank test results will exceed the limit of blank and will be erroneously considered detected. This probability is small usually only five percent. Below the limit of blank assays cannot distinguish signal from the level of noise because of low analyte concentrations. The limit of detection is typically defined to estimate the concentration where usually five percent of the test results shown on the graph in pink were erroneously considered not detected or below the limit of blank. This is usually established by measuring samples with low analyte concentrations more than 60 times with at least two reagent lots. For assays test results below the limit of detection indicate that the analyte is detected but are not analytically reliable, that is precision is not acceptable.

Meanwhile the limit of quantitation is the lowest amount of the analyte that could be reliably detected and at which a stated performance goal is met.

For tests where there are no recognized standards available such as troponin assays the performance goal is based on acceptable imprecision.

I would like to highlight some challenges we have observed regarding detection limits during review of submissions. The first is that we have observed different detection limits for different reagent lots. The following table shows hypothetical data for the LoD, LoQ and troponin concentration at which the percent CV is less than ten percent for a hypothetical troponin assay. Each parameter was determined on three different reagent lots. As shown in the table the limit of detection varied from 0.8 to 1.7. The LoQ ranged from 1.2 for lot one to 2.5 for lot three. Finally the troponin concentration with the percent CV less than ten percent ranged from 4.3 to as high as 8.9. These data highlight that the manufacturing standards are not able to produce consistent detection limits for different reagent lots.

Second this table shows the calculations for percent total error relative to the target value for a hypothetical troponin assay. The percent total error

is calculated based on measured troponin value relative to a target value from two different reagent lots. As demonstrated in the table the percent total error is very high at the low troponin values. Thus a value of one nanogram per liter may not actually be one nanogram per liter. Total error decreases as you get closer to troponin values in the clinically relevant range. These data table also show that total error can vary significantly by lot, for example, 80% versus 140%, again highlighting that the manufacturing processes for these assays cannot ensure consistent performance of troponin values at low analyte levels.

Finally other challenges of using these values clinically include that calibrator target values are often not near the LoQ and thus these assays are not sufficiently anchored at the low end of the analytical range. Control target values may also not be near the LoQ and thus assay performance may not be sufficiently monitored over time.

Another point of consideration is that FDA has historically reviewed the clinical performance of troponin values at the assay cut-offs and clinical

performance has not been historically assessed at the LoQ for troponin assays unless the cut-off is near the LoQ.

Moving on to some trial specific issues. Sponsors often perform clinical studies at multiple clinical sites. Sponsors then sometimes choose to perform testing of clinical specimens at multiple laboratory sites. We sometimes observe different results depending on the location of the testing site. The following data table shows a clinical performance estimate from three different labs. The sensitivity ranges from 74% at site two to 92% at site one. Similarly the positive predictive value demonstrated different values depending on the particular laboratory. Such data raised questions from a statistical issue on whether such data are poolable. We have also observed different analytical and clinical performance using the same assay reagent on different analyzer modules from the same manufacturer even when these analyzers are considered family members. A goal of conducting clinical studies for troponin assays is to generate the data that intended users can expect to

see in their lab.

Finally another trial issue that we see is that of when to perform and accept repeat testing. Troponin is typically run one time and reported to users in clinical laboratories. Repeat testing is not usually performed unless the QC is out, there's an analyzer alarm, a delta check or a clinical issue that arises that prompts the clinician to request repeat testing. Occasionally sponsors will do retrospective analysis of study results already performed and will identify QC issues such as drift or trend, an increase in instrument alarms, or some other issue and decide to repeat a subset or all of the samples. A potential issue with this approach is that it does not affect how troponin testing is performed in the intended use population where testing is usually performed only once. The goal of the analytical and clinical validation testing is to generate representative results for troponin assays that intended users will observe and repeat testing of a subset of samples may result in bias.

The final issue I will discuss in this

presentation is the emerging issue of biotin supplementation and interference with laboratory test results. There are medical conditions such as secondary progressive multiple sclerosis in which high doses of biotins may have a clinical indication. However biotin is being increasingly marketed as a beauty supplement for hair, skin and nails. Tablets sold over the counter contain ten milligrams or higher. While the actual amounts of biotin being taken by consumers is unknown there are increasing reports that consumers are taking biotin in amounts well above the daily recommended intake which is thirty micrograms. A potential medical issue with these high doses of biotin is that biotin in a patient sample could interfere with a broad range of diagnostic tests specifically assays that use biotin streptavidin binding. Biotin streptavidin is a popular component of assay architecture due to streptavidin's high affinity for biotin in binding under a wide variety of chemical conditions. Unfortunately susceptibility to biotin interference is variable in magnitude and can cause either falsely high or falsely low results depending on

the assay design and conditions. A proposed strategy for some tests for avoiding biotin interference is to have the patient stop biotin and wait several hours to several days before laboratory testing. However, this is not an option for troponin assays for a medical emergency such as acute myocardial infarction. Additionally patients may not tell a physician they are taking biotin or may not know that they are taking biotin. And health care providers may not always inquire about biotin intake.

There is also not a way for a lab to easily identify specimens with high levels of biotin that may cause interference with troponin assays and labs may not have an alternate method of readily available to test patient specimens.

FDA has received at least one death report associate with biotin intake that was associated with a false negative troponin result. The clinicians in this case did not find out the patient was taking biotin until several days later. This case highlights issues with high doses of biotin and interference with troponin assays.

Released today we have a safety communication on biotin interference. You can find this safety communication at the link listed on this slide.

I'm going to conclude with our list of questions for the panel. Our questions are: How can specimen stability be incorporated into the trial design? What differences in troponin results have intended users observed with different tube types? How can manufacturers better control test performance at the limit of quantitation? How can the reliability of troponin testing at different sites be maintained? And how could we prevent unsafe use of troponin devices because of biotin interference? And what steps are necessary to address this issue?

Now I'd like to go ahead and turn this session over to our panel for discussion.

Brittany Schuck will be monitoring the session and we will have the panel introduce themselves.

Pre-Analytical and Analytical Considerations for Clinical Trials

DR. APPLE: The same Fred Apple.

DR. CHRISTENSON: Rob Christenson, University

of Maryland, School of Medicine.

DR. GREENE: Dina Greene, University of Washington.

DR. PHILLIPS: Jane Phillips. I'm from Roche Diagnostics in Regulatory Submissions.

DR. SAENGER: Amy Saenger, University of Minnesota.

DR. CAPOSINO: I'm still the same Paula Caposino.

DR. WELSH: Hi, I'm Kerry Welsh, I'm a reviewer in the Division of Chemistry and Toxicology Devices.

DR. SCHUCK: All right. I want to reiterate what people have said previously and thank the panelists for joining us today for this discussion topic.

And I would like to start off with biotin interference as Kerry mentioned we did publish today a safety communication on biotin interference with laboratory assays. And so we are looking forward to hearing feedback from the panelist and the participants on how best to address this issue and what ways we may

address issues with biotin interference particularly for troponin where waiting would not be an option.

DR. APPLE: So just to go down the line I think it is obligatory on the manufacturer to identify whether or not they have biotin interference especially around the 99th percentile because we know that is the microscope that is being looked at. Remember we can't lose track if there is still a clinical component to this and the clinicians have to look at it and we're looking at a change, not just one point in time as we heard I think Frank said you just can't make a decision on one point often especially diagnostically. So I think we have to be able to pick up a package insert, read what the interference is, positive or negative, and that is something we, as laboratorians, have to make sure our clinicians understand whether or not they know their patient is on that or not and go after the concept of looking at a moving increasing or decreasing troponin or a static increase of chronic disease, not an acute disease.

DR. CHRISTENSON: As Fred said we in the laboratory consider the instructions for use or package

insert to be you know sort of a book of truth what the FDA has examined. I think biotin should be handled like any other interference. Certainly at the -- so all interferences at the 99th percentile or at the cut point that is used would be very important. Also it has to be determined whether biotin that you could adulterate a sample with out of a vitamin bottle is the same as the biotin interference that a patient who is taking biotin would have. So is that important?

I feel certain that my colleagues know more about this matter than I do.

DR. GREENE: First I'd like to thank Dr. Welsh because I think that was a really excellent introduction to a lot of the things that we do think about all of the time. I agree with Dr. Christenson that the biotin -- so not being aware of the specific case that's being talked about or having read the bulletin interferences happen all the time in the laboratory. We depend on our clinicians to call us and say we think something is wrong with this specimen. And then we work it up appropriately. I again am not familiar with this case and the only and I'm probably

going to get in trouble for saying this but one of the main times that I think CKMB is appropriate is when we think that there's a troponin interference because at least then you can kind of look at the two and say whether or not it does look like there is cardiac ischemia occurring. I think that -- yeah, I just agree that interferences happen all the time. And this is a tragic case but we have to be willing to just work up samples and to investigate further to see how these are going to influence results.

And definitely any manufacturer that is using biotin as a reagent needs to perform the studies. It doesn't hurt anyone that is healthy to be taking biotin supplements. It is easy. I would participate in a study in a second where I had to take biotin supplements and allow for samples to be drawn. So I think that those things can be very easily accomplished either biologically or by adulterating samples with purified biotin.

DR. SAENGER: And I'll just add to Dina and everyone's comments that really we don't actually know the frequency at which this occurs or that even

patients who are presenting to the ED how often they are taking biotin. I think there are some studies in the works. But I guess I haven't seen anything that's been formally published quite yet. But to me it is another interference; it can be a problem. So can heterophile antibodies, to me bigger issues are probably hemolyzed samples and how instruments are detecting it or not detecting it because that is a huge patient safety issue. So I think it is another thing that we have to be aware of and cognizant. But I think until we know exactly how big of a problem it is specifically for troponin I guess I personally wouldn't get too worked up about it just yet.

DR. PHILLIPS: And to Amy's point I want to just say from a manufacturer's perspective this is an evolving issue. So it is pretty much new to us. We thought okay, biotin interference we report in our product insert. So just the information on who is taking how much and what is the consumption versus the purchase of biotin. These are things that need to be looked into by manufacturers. We have done a pharmacokinetic study that has recently been accepted

for publication. I can get that reference if people want to see it. But what we do see is the biotin thankfully is cleared relatively rapidly. So not 20 hours later but more like within three hours. So that's at least one good piece of information that we will provide to the labs. I think there is a lot of education that needs to be done on the point of the manufacturers not only to the laboratory but also to the end users.

DR. SCHUCK: So I would follow that up with one proposal we have heard is to label with waiting or recommending time after last consumption for biotin testing. Is that information helpful? You mentioned a lot about including this in the labeling and knowing the levels of biotin interference. But what additional information may be helpful?

DR. CHRISTENSON: I think that would be useful because FDA I'm sure realizes that it is not just troponin, it is anything else that was being -- any kind of endocrine test that were being ordered on these samples. So probably what labs ought to do is just put out sort of an alert to all the sections or have some

note go in that this patient's on biotin. Is the pharmacokinetics of biotin is that very well elucidated?

DR. PHILLIPS: I don't think it is. I think biotin was just in your multivitamin up until the hair, skin and nails craze. So I think it is really like I said an evolving topic. And I think more pharmacokinetic data need to be generated.

DR. SCHUCK: Thank you. Let's go ahead and take a comment from the audience.

DR. JAFFE: I can provide a little bit of information. We've developed a Mass Spec biotin assay and have been looking at patients and we do see -- and we originally appreciated this because of its effect on TSH predominantly. We know that it affects NT-pro. We did not see in our institution at least an effect on the fourth generation troponin T assay. We do see an effect on the fifth-generation assay if we use a cut-off of the change of the lowering of three nanograms per liter. That occurs at a value of about ten with our mass spec assay recognizing it only gets intact biotin and not all the fragments. And that is about

two and half percent of our ED population. So it is potentially an important issue for every manufacturer to look at closely.

The other thing that we do know is that biotin is cleared renally so that renal failure patients are at particular risk. I hope your alert covered that aspect as well.

DR. SCHUCK: Thank you for that. And I do what to circle back to frequencies. So although we don't have a lot of published data on frequency, MDR analysis and complaint analysis would indicate that frequency is increasing particularly with the concentration or the dose or the level of biotin being ingested. So maybe you can speak to that a little bit further.

DR. GREENE: I'm just curious is there a method to label the biotin supplementation?

DR. SCHUCK: Not that we are aware of. Yeah, not from our side. And so I'm curious about additional ways to address this issue beyond labeling particularly for troponin assays because labeling for troponin assays is going to be very difficult given the time

sensitive nature of troponin.

DR. CAPOSINO: I'm curious as to when somebody suspects an incorrect result for troponin. For example if there is no other, you know there are people who don't have anything other than an elevated troponin. Where does that -- where do you think that comes from where somebody would question a negative result and how often are backups, different biomarkers being used? I think it would be interesting for us to hear how a clinician would suspect that incorrect test result and what other information you may have as the lab director to think about that that it could be wrong?

DR. CHRISTENSON: So often times -- I'm sorry, Amy. Often times what we see is there'll be a dramatic change in a value. So you had ruled out MI because the two-hour sample was negative and yet you measured at two and then for some reason you order a six for completeness and there it is. And with rapid clearance I mean that's the kind of thing that might happen and a clinician would call and say well this -- what's wrong with your test. We had ruled out but now there is a diagnostic value for troponin and nothing has changed

clinically. So either one way it is a positive or as a negative would be one way that we would be called.

Amy?

DR. SAENGER: Yeah, I mean I would say it is a clinical suspicion. We wouldn't necessarily know in the lab but I would say overall it happens infrequently. But generally we have ways to investigate either false positive or false negatives. I'd rather use a different, if we use I I'd rather use a different I than CKMB, no offense. And I mean if it is a heterophile, we have heterophile blocking tubes or we do dilution. I mean there are general routine things that we do. For biotin it is a little bit different because there is not a lot of ways to investigate it other than kind of asking the patient the specific question which they may or may not know the answer to.

DR. SCHUCK: And I'm also curious how frequent it is that your secondary method or your additional analyte would not be a biotin test. So a lot of these platforms have biotin for all of their analytes. So what happens in that circumstance when you don't have a

backup that is not biotin?

DR. SAENGER: I think we'd want to have a backup that wasn't -- didn't have the interference for all immunoassays. I mean it would be advantageous to do that not just for troponin but anytime there is a question.

DR. APPLE: So we have sent out a memo to our medical staff. So it has been part of the medical exec that people know that here are assays that have biotin problems. So that is about as best as we can do. When we have a flyer whether it is a sodium or a troponin or endocrine test you can't figure it out with your clinician sometimes, very rarely, but sometimes we'll send it across the river to another hospital and we'll pick an assay as you said that doesn't have that known interference and I'd say 50% of the time it comes back the same and 50% of the time it doesn't. And what does that mean. It is really more of an academic question. But clinically I think we run across this all the time and I'm drifting here to another question but we avoid Lit Hep specimens because of the flyers that we used to get. So you maybe have a Lit Hep that causes and the

biotin causes a decrease and you come out normal, what does that mean. I am being facetious but we have these problems. This is our business. So we have to deal with these kind of things on a day-to-day basis. It is not just like the package insert says it is perfect. We have all kind of issues that we have to deal with like this. Troponin is no different.

DR. GREENE: Different assays are influenced to a different extent and even some that use biotin aren't necessarily influenced. And so understanding the bias that occurs between the different assays what would be the expected bias and did that just follow that trend can be helpful. But it is very difficult because yes, I can send the sample to a hospital down the way but that's not going to give me a very immediate result.

DR. SCHUCK: All right. Thank you. We'll take one more question on this from the audience and then I'd like to move on to some of the other topics. But this is definitely a topic we're interested in and we'll circle back to.

DR. JAFFE: It is not a question. It's really

a comment. I think that the idea that the lab can pick up everything is not really something that we can expect laboratories to do. We can debiotinylate samples as a way of doing quality control but it is a painful and difficult method and we can't identify them upfront. It is why it is so important that the clinical link to the interpretation of all of these assays is what is so important because it is the clinician who says this is a high-risk patient with a low troponin; what's going on? And unfortunately one of the things that happens as a combination of approval and marketing is the idea that okay, these things just roll downhill. Value below the LoD couldn't be anything but a rule out even if the patient has chest pain and ECG changes. And I'm being hyperbolous a bit but I think we have to be very careful to make sure the link between the clinical and the laboratory is tighter than it has been heretofore.

I'll say one other thing and that is we in the lab have tried to surveil what's going on. So for example there was a craze that went on starting to use biotin in patients who had MS which we saw in the

literature and then talked to our neurology colleagues to sensitize them because they were starting to see patients who were coming in on therapeutic and much higher doses of biotin because some people think it helps certain subsets of MS.

DR. SCHUCK: Thank you for that comment. I am going to switch gears here a little bit --

DR. CAPOSINO: Can I just ask --

DR. SHUCK: Oh, just kidding.

DR. CAPOSINO: Just one thing is this something that is widely known among clinicians that biotin could -- okay. So hopefully you guys will look at our safety notice.

DR. SANDOVAL: A quick comment on that. So I've rolled this out at seven of our hospitals so we have 250 physicians and you know maybe another 150 physician assistants. Several of these people are at different stages in their career, some people are just learning about medicine. Some people are at the tail end. Some people in the middle are just worried about getting their kids home from school. And you throw out something on the report. They'll see it the first time

and after a thousand other troponins they'll forget that it exists. You put it on the package label. I didn't know there was an FDA label on any blood test that I've ever ordered until about a year ago. So think about how most physician look at these things. And our ability to educate people is a short period of their attention span. So this is going to be a bit of a problem as far as how we generate this message and keep it in people's minds.

DR. SCHUCK: Thank you for that comment. I will switch this boat. One other comment. Is if the clinicians are educated or are more broadly educated on the issue of biotin interference and they are to put that information in the lab report that comes to you; do you have the opportunity or the tests available that don't use biotin so that if you did have that information that you could make a different choice for your test?

DR. GREENE: There have been pretreatment with streptavidin coated beads published in the literature that does show that you can use the recovery results clinically but that would have to be validated by each

individual laboratory if they didn't have an alternative platform which most of us don't.

DR. SCHUCK: Any other comments from the laboratorians? Okay. Thanks.

We will switch over and Dr. Apple you mentioned differences in different tube types and that was something we were interested from Dr. Welsh's talk in discussing. Do you want to elaborate on the differences you see in the laboratory? And anyone else?

DR. APPLE: So a couple -- do you want to stick to just the tube types? Nothing else?

DR. SCHUCK: For now. But if you have additional comments within the topic that's fine.

DR. APPLE: Like three or four points. So I mean a couple things. We live in a world that you have to make a decision on what you are going to measure serum, lit hep plasma, EDTA plasma. So as a clinician -- as a laboratorian I've noticed over the years of troponin that we always use lit hep and we got a lot of flyers. So we had a policy of repeating the first positive every time we got a troponin positive. We

since have validated in the U.S. in our own laboratory EDTA. And I think you commented on it that EDTA looks like to me the troponin assays I play with the cleanest samples that I measure. How do I determine a clean sample? Chris deFilippi, Allan Jaffe, none of my clinicians, Bob McCord, no one calls me about a flying -- a result that is odd. So we have -- it is not FDA cleared in the U.S. but we have validated and have been CAP inspected and shown them all our validation data. It has worked incredibly well. So I think you have to find your sample that is going to work for your system. And agree if you are going to have a claim for a lit hep and a claim for EDTA they have to show some type of studies that show that they both work.

The other thing that you pointed out was some of your fictitious or hypothetical scenarios down at the low end of that high sensitivity assay. A lot of those are below the 99th percentile and a lot of those are probably within analytical imprecision acceptance. I was doing some calculations. There was 25% variation between a couple of your numbers that were different. And I would say I wouldn't even lose any sleep over

those because we know that we measure a sample three times that you are going to see variation within those sample type or between sample types that will give you that variation just because of the lot number differences that we use or the calibrators or the reagents. So I kind of view that and Rob and I participated in the study, was looking at simulation and samples and imprecision even between sample types up to 25% had no effect on miscalculation or misidentification of a reference sample being abnormal or normal or on even clinical outcomes. So I think that leaves and we'll get to the LoD and I won't -- I'll wait until that question.

So specimen types, imprecision I wasn't really concerned that much in what I saw there, no different than potassium when I look at sample types and look at instruments we see variations like that.

DR. CHRISTENSON: Many of the problems with specimen types especially in either EDTA or heparin is because you know it takes eight, according to the package insert for tubes it takes like eight inversions in order to mix it properly and what happens clinically

they draw the blood, they put it on a table, then they put it in a tube and they send it up. There is not a lot of inversion. So I think we need to remind our clinical staff about that if they see these issues. I agreed that EDTA seems to be a much cleaner sample than heparin but a lot of times because heparin seems to be that the clotting it is kind of ongoing still when you look at a lot of specimens. So I think that is the micro clot thing where EDTA doesn't seem to have that problem. Heparin I mean serum would be a wonderful sample as well but the issue is that you have to wait at least 30 minutes and what if the patient is getting heparin which most of these patients may well be getting heparin, then your serum is going to clot slower. So it is an hour before you could probably do the assay. So we are kind of stuck with doing -- and I think the bottom line of it all is we have to just demonstrate I mean here's something where we don't have to rely on opinion, we can actually get data to show. It is sort of like what Fred was talking about that he showed his CAP inspectors that it was validated using EDTA.

DR. SAENGER: And I'll just add to that my prior experience at Mayo we saw a lot of -- not a lot but fairly frequent flyers enough that it caused some patient safety issues. So we actually validated the rapid clot serum tube and granted it is more expensive but I think serum can be a cleaner sample and lithium heparin is generally considered a dirtier sample. But the benefit to that is might have been more expensive for the actual tube but visually it looked different and so sometimes that helped facilitate also a more rapid turnaround time. And it clotted I mean very quickly.

DR. GREENE: Unfortunately I have to contradict Dr. Saenger because we use the rapid serum tubes in my laboratory, we still see a significant amount of flyers where we have -- I agree that it is our assay but even irrespective of that we can't put those on our automation line because they are these like massive clots that we'll get in the tube that these samples need to be treated separately or they will clog up our chemistry analyzers.

DR. APPLE: So as Dina said our automation

line. So we have a dedicated draw for troponin because we have to worry about turnaround times because of the pressures from our emergency department. So therefore we can get around putting it on our automated sampler, we have a sample that goes right to the station where troponin is done and that is how we keep our turnaround times 45 minutes or less. But if you are stuck putting it on a front-end automation you are going to run into more problems especially with troponin. And depending on the specimen type and sample type.

DR. SCHUCK: Okay. If there is no more comments on that particular topic I would -- oh Paula?

MS. CAPOSINO: I would just make a comment. I think with some of these differences when we note them as observations it is when it's outside of the expected imprecision. So I would just like to make that comment that we understand the imprecision of the assay and we consider that. So in these cases where it was really showing a change above and beyond what would be expected because we know that if you run something twice you are not going to get the same exact result. So we are mindful of that. We work with our

statisticians to understand what could we expect with the imprecision. So I just wanted to make that comment.

DR. SCHUCK: Thanks, Paula. So with about ten minutes left I want to turn to sample stabilities. So we heard this come up in earlier sessions particularly with clinical trial design. We would like to receive comments from the panelists and anyone in the audience regarding sample stability, of sample type, of time of condition of processing and handling of these samples.

DR. APPLE: Since we are short on time I need to just talk about LoD for a second; is that all right? Because that is really a key point.

DR. SCHUCK: Yes, yes.

DR. APPLE: And I'll talk about sample stability. So we heard the discussion with our clinicians before the power of LoD. And you can look at 14, 15 international studies done, trials that have shown, we talked about a negative below the LoD and looked at a negative EKG and the power of being able to say a low risk patient with less than point five percent bad outcomes in 30 days. You can make clinical

decisions.

So I know that the -- has developed some assays that have been cleared that the 20% CV value has become somewhat, I'm not sure it's a FDA requirement. I'm not sure it's a manufacturer necessity. But if we don't allow -- if the FDA doesn't allow if we are going to this rule out concept which I think is going to be very powerful for financial savings for our emergency departments and our patient care to report results down to the LoD. So a 25% imprecision LoD what is the difference. The data is overwhelming that it'll show that we can make good -- my clinical colleagues can make decision whether they want to send home a low risk patient with a normal EKG and a low value. So the data we did in our U.S. trial and other European trials is I recommend look at this meta-analysis that was just published in JAMA last week, Andy Chapman is the first author out of Scotland. Looked at over I think 20,000 patients from multiple studies, the power of the LoD or the lowest concentration that gives you about a 99% sensitivity and a 99.5% negative predictive value you can send anywhere from 15 to 40 plus percent of

patients home and if we are going to limit that result it has to have a 20% CV or less, we are going to miss out on that opportunity for good patient care and patient management in my opinion.

DR. LIAS: So Fred can I weigh in here. This is Courtney Lias from FDA. My colleagues at FDA may be very clear on this point but I'm actually a little bit confused and one of the things that I want to make sure we understand today is that we are crystal clear on why people are requesting values down to what you are calling the LoD. So some of the things you are saying make me think that maybe you are not talking about the LoD. And that maybe the terminology that we are using is a little bit different. And but it may not be that so I just want to make sure I understand. So if you mean that you would rather have a different definition of the acceptable imprecision of the assay to define quantitation that is different than saying you want to report down to the LoD. But if that is not what you are saying then I think that we don't necessarily understand exactly what it is that you want, so that we can help you get there.

And also it's clear it sounds like it does relate to the issue of rule out which we've already discussed we are completely open to and anyone can design their assay as a rule out assay and tell us how they've designed it and look at how it works. It is also okay to specify a negative predictive value as your cut-off before trial and do it that way. So all of these things may be possible but at the moment I am not exactly clear I understand what you mean by we want to report down to the LoD?

DR. APPLE: I will be corrected if I misspeak. The LoD, the world's expert opinions have opined that at the 99th percentile you should have an imprecision at least ten percent if you call it a high sensitivity assay and not have an assay reportable if it's greater than 20% at the 99th percentile where you are making a diagnostic decision. For the LoD is a rule out decision and it is not -- it is independent of what the imprecision is down there. So it is not a new definition for any MI even diagnosis, it is going down to the LoD could be 30% imprecision or 25%. The data speaks for itself that it is a very powerful tool. So

there are two different issues. Does that make sense? There is a rule out --

DR. LIAS: I think it would be helpful for us to understand what it is somebody wants out of a rule out assay because it isn't clear to us from the data we've seen that just going down to the LoD would result in the type of performance that you're talking about. So it sounds like one you just want to be able to have rule out assays that work. That maybe you'd be willing to accept a little bit more uncertainty at the low end to get that. I think we should stop using the term LoD because I'm not sure we're talking about the same thing.

DR. APPLE: But we are now. We are just the way you defined this, the way it was defined. Allen maybe you can comment.

DR. GREEN: Well, I think this comes down to the difference between again the clinical and the analytical. And so no, I don't think that at the LoD we have the analytical performance to really accurately quantify the assay at the precision that you want. But clinically what's been coming out is that this value is

very, very important so Pete Kasak just published a paper, I think it just went to press today about the variability at the LoB and the LoD of assays and yes, they are variable as you showed. However when you do meta-analyses as Nick Mills' group just showed and you are looking at people that aren't focusing on the analytical as much but really focusing on the clinical this value gives you what you are looking for.

DR. LIAS: It is still an issue of conversation because the LoQ is defined as you need it to be defined. And/or you define that assay as a qualitative assay of some sort. So I think that it is important that we decide what's clinically needed. So I'm telling you I'm agreeing with you.

So I want to clear up the idea that we're against reporting something that is clinically useful. So if the clinical community makes a definition that this is the information that is needed, test designed in this manner that performed in this manner would be helpful then that is not an issue for us. So at the moment that isn't what is happening. So it is helpful to hear the discussion that the clinical community

wants tests that can give some information at the low end. It would be even more helpful for us to understand a little bit exactly what that information is.

DR. CHRISTENSON: So fortunately at this side of things on this topic LoQ is the one that is a little shaky because what is it the reliability, the concentration at which the results are reliable; right. Well, humans make that decision. It is not made by mathematics; right. LoD is very well defined as we saw. So what --

DR. LIAS: But the LoQ and the LoD could be the same. It --

DR. CHRISTENSON: What's that?

DR. LIAS: You define the performance parameters around an LoQ for what you need. So the 20% doesn't come from FDA.

DR. CHRISTENSON: Yeah.

DR. LIAS: The 20% I think is out of the air and people just commonly use it.

DR. APPLE: It's from the TSA world.

DR. LIAS: It could be something different and

it could match the LoD. So that happens in some assays where the LoD and the LoQ end up being the same.

DR. CHRISTENSON: But risk stratification also down to the LoD there is abundant literature that shows that that is valuable as well. The LoD as defined like what you all did just a few minutes ago. So what we need to do is be able to make sure that that remains reliable, that that remains constant. You showed numbers some of which were within that 95% confidence whatever but the point was that there's heterogeneous results in there and we need to -- the manufacturers I think just need to make sure that those from lot to lot that it’s consistent.

DR. LIAS: Yeah, I don't know that it is widely recognized how much variability there is. And I want to emphasize that I am not arguing. Actually this is very helpful to understand that we need to use language that gets us on the same page because I think some of these misunderstandings is simply that sure one can define a clinically useful level or test design any way you'd like and that just hasn't been how people have used these terms. So thank you for clarifying that

because it is really helpful.

DR. GREENE: And also I just would like to add that the manuscript that Fred alluded to earlier that should be going into press in Clinical Chemistry we do recommend that just like folks check their analytical measuring range every six months that they also do a check of their limit of the blank and limit of detection to make sure that their assay has stayed stable. And when they are experiencing instrument malfunctions that that is something they check just like they check QC and calibration.

DR. SAENGER: But the problem is in the U.S. we won't be able to do that, you know, so that will be outside the U.S.; right? You will only be able to check LoQ.

DR. GREEN: Oh, yes, I guess it depends on what you can get off the instrument. Currently I can get whatever values I want.

DR. LIAS: If assay manufacturers talk about the design of their assay and what they are supposed to be doing with it and if the clinical community is clear on what they need and what definition of clinical

meaningful values would be then the manufacturers could use that as support. So none of this is in conflict. At the moment that is just not what's happened.

DR. JAFFE: I'll just speak to the clinical circumstance. There's sometimes when the data is there big ends and it obscures subsets but in this instance I think it is pretty clear they're for both troponin T assay and for the high sensitivity Abbott assay the data were very complete that using values and the magic number has varied in fairness so there is some ambiguity. Some people use the LoB originally. Some people use the LoD. Some -- Nick Mills originally used five. Then there is a meta-analysis that shows the data. It really doesn't matter because empirically you get down to the LoD the data are so overwhelming that clinically it is one and done for low risk patients. And my biggest fear about advocating it is that it is going to become something that is used in non-low risk patients. And I think that would be in error.

DR. APPLE: Courtney I'm going to give you a quiz later about if you --

DR. SAENGER: I think the other thing to be

cognizant of is the whole IFCC definition of the percent of detectible males and females is also based around the LoD. And so in the U.S. we are a little bit hindered now at least with the Gen 5 troponin T to even do those kind of studies if you wanted to. We'd have to do a research process which -- or send it outside the U.S.

DR. CHRISTENSON: I guess I would ask FDA what is the issue about reporting down to the LoB. If we could make the LoB -- or LoD I should say, LoD is what I mean, LoD stable what would be the problem with reporting that to caregivers?

DR. CAPOSINO: You know I think we would want to understand the analytical validity behind that.

DR. CHRISTENSON: So that is the issue is that the analytical validity has not been well established down at the LoB.

DR. CAPOSINO: I think as Kerry explained we are looking at where the test result is clinically meaningful which is sometimes folds above that. So we are not looking closely at study design. You know we are not making sure that the study is designed to

actually show an LoD and if we understand that that is what the device needs to do our review would be a little different.

DR. APPLE: So you are saying that if a manufacturer comes in with their clinical sensitivity set specificity diagnostic data they show what their -- what we consider an LoQ is 20% CV but then they show you clinical data down to the LoD which might have a 30% CV but the data is clinically relevant that is something the FDA would consider?

DR. CAPOSINO: I don't think we are opposed. We would want to see analytical validity at that claim to make sure that if you are making a decision on a number three that the next lot that number isn't eight. So that's --

DR. SAENGER: But I would think most of the clinical trials that are ongoing right now looking at sensitivity specificity are using the values down to the LoD. They are not censoring them to the LoQ. So the clinical data that you are getting is based off LoD I'm pretty certain.

DR. APPLE: So it is up to the manufacturer is

what you are saying is to show the data and provide the evidence of the clinical indication.

DR. GREENE: Can I ask a clarifying question? Does the FDA prohibit manufacturers from allowing the lab to see those results even because I know that there are certain manufacturers at this point where you can see every value down to the lowest of the low and other manufacturers that censor those values. And so I'm wondering is that a manufacturer choice or is that something that the FDA is limiting?

DR. CAPOSINO: So the reportable range is defined by the manufacturer and that is the range that is analytically supported. So outside of that range the test result may not be reliable or has not been shown to be reliable. So that is what is reported.

DR. APPLE: The question I think to -- this is a research question. Let's say you are only reporting down to the analytical but you want to do research like Amy suggested, we'll pick on Roche as an example, they report out less than six. Let's say we want to do research and measure down to the three LoD. Is it the manufacturer's decision they can't report to us for

research or is it the FDA said you can't report that even though it is not patient care oriented? Who makes that decision? Because that's been a question not just for Roche but for many assays over the years.

DR. CAPOSINO: So we review what is used for clinical use. So we -- we're not involved in the research use. I'm not sure --

DR. GREENE: It is a very practical question. So when I look at the LoB for every assay that I have so I can make sure that the clinical values that I'm reporting are at least you know three to four SD away from what the LoB of the instrument is giving me. I do that for creatinine or alcohol or any test that I measure. And so to not be able to see those values as somebody that is being relied on to report out these results is very difficult. And so that's I think not even for a research purpose but a practical purpose of maintaining quality in the laboratory that these values are very important for us to see.

DR. PHILLIPS: And I think in general to Paula's point the manufacturers have to prove that there is a use for those values from a clinical trial

perspective. And so we have not done that yet. If we approach FDA in the future with a risk prediction claim then we might have a totally different discussion.

DR. APPLE: Just one more comment. So Jane, I don't want to be argumentative but in your case that is not accurate because you have European, the same data, the same assay in Europe where they use of the LoD has shown considerable clinical value and for the U.S. population to be able to study the assay to get data we don't have the opportunity. So that is a point of discussion I think with the company what I heard, not with the FDA.

DR. PHILLIPS: Yeah. And I think that is something we have to work on more collaboration within the U.S. to generate the data.

DR. SCHUCK: We are five minutes over time but we'll take four more minutes to answer a couple of -- or have a couple of comments from the audience.

DR. GUTIERREZ: Alberto Gutierrez. I'm actually ex-FDA so what I say is only representing me. But I can give some perspective here. And you are all caught a little bit in what's the chicken, the egg.

You have to remember that when these analyzers are cleared they go not only to high complexity laboratories but moderate complexity laboratories, many times, and so what the instrument manufacturer does obviously is make their instrument so they can be used. Now, there are opportunities I'm sure the agency would be open to in other areas that has been done and there are opportunities in which you are able to either through software release the results so that you can do things that a laboratory would need to do but you would have to approach the -- that I think needs to be worked out between the agency and the manufacturers and try to get something that is useful for the clinical community overall. But it is a good question. I think it is just a matter of it’s happened that way partly because the data that the agency gets and what instruments are going to be used for and so the labeling it just falls that way. It is something that I think needs to be worked out between the clinicians and the FDA as to what is useful and why you know sometimes what could be done to make it more useful particularly for those clinicians, for those laboratorians that need to get

data so that they can run the laboratories appropriately.

DR. SANDOVAL: Let me just make a quick comment. I can help with a little bit of confusion. So I think several people have made these comments but I can help some with degree of confusion so the 99th percentile and I think Dr. Jaffe made this very clear in regard to the universal definition of MI still the threshold that we are going to use that there needs to be one concentration above the 99th percentile with the rise and fall to diagnose, to rule in myocardial infarction, that is still the standard, that is what the clinical practice guidelines say. However for those of us that have to follow the large burden of literature publications what I think Frank Peacock said earlier and it applies to this is that a lot of the emerging literature has focused on low concentrations that are not the 99th percentile. They have been across a range of concentrations. There are some studies that have looked at LoB, some at LoD. There are even some from Pete Kasak, I recently I think in Clinical Biochem looking at LoQ. And the largest ones also making those

meta-analysis looking at concentrations that are not analytical thresholds, they are thresholds that were developed on the basis to meet a clinical need such as the five nanograms per liter. That is not the LoD. So the point is right now the indications in the insert packages are to aid in the clinical diagnosis of acute myocardial infarction and I think Frank alluded to this earlier that right now aid in the clinical diagnosis of myocardial infarction are mostly rule in a sort of phrase. So right now we can say you use the 99th percentile rise and a fall and certainly it is an aid to rule in myocardial infarction but you can use many of these low concentrations such as LoD or other concentrations that are not analytical such as five -- identify patients at low risk that can be discharged quickly. And that is regardless I also want to make a point that the LoD across different manufacturers some of them as I think Amy alluded just measure 50, 60 different thresholds but there are some that measure well over 90% of all normal individuals. So how would it be useful for assays that are super extremely analytically sensitive to have 95% of values measured.

So my point is for in the ED it wouldn't be as helpful because no one is going to -- there is a range, you take a step back, clinically it’s a little bit more than analytical threshold, which is the threshold that allows the largest proportion of patients to be identified as low risk that can go home regardless of where that concentration is. So there are two issues 99th percentile to rule in. And a new movement to identify this low concentration in low risk populations to send these patients home.

DR. SHUCK: Thank you for that.

This has been a really great conversation on detection limits among other things. We are out of time for the session. I would encourage folks to comment during the public comment session because we did not get to sample stability during this panel. But I believe it is now time for lunch, a break and we'll be back here at 1:00.

LUNCH

DR. LESSARD: I think we are going to get started with the next session. If you will please take your seats and if I could ask the panel members to come

up and take your seats. Thank you.

All right. Welcome back everyone. Just before we get started one quick announcement. Right now there is one person signed up for the public comment session. If we are missing anybody please let us know, please sign up. We'd love to hear your comments.

So welcome to our next session on Clinical Trials for Point of Care Troponin Devices. My name is Juliane Lessard. I am a reviewer in the Division of Chemistry and Toxicology Devices in the Office of In-vitro Diagnostics and Radiological Health.

Before we launch into our panel discussion on this topic I'd like to briefly touch on the benefits and challenges of using troponin devices in a point of care environment and then talk about some considerations for designing clinical trials to support the performance of point of care troponin devices based on the examples of different challenges that FDA has observed over the years.

FDA understands that there are unique benefits and challenges for point of care testing for troponin

device and takes these into account during pre-market review of point of care troponin devices. Some of those benefits include the convenience of on-site testing, whole blood matrices that require less processing and real-time availability of test results. Research indicates that the immediate availability of point of care test results can help lead to more timely intervention. And in some cases for example on free standing emergency rooms a point of care device may be the only device available and it may not be feasible for physicians to wait for a troponin test result from a central laboratory. Point of care settings where troponin devices are used are typically very busy and less well controlled for environmental factors such as temperature. Point of care operators often multi task testing and patient care and they typically have less training on how to perform invitro diagnostic tests compared to clinical laboratory professionals.

Each troponin device including those used in point of care settings has its unique performance characteristics and limitations. Results from one assay are not typically interchangeable with other

methods. For example serial testing of a patient with different point of care devices or then a laboratory method may not provide the most accurate clinical picture.

When reviewing troponin devices FDA considers the full scope of the device including access to testing, turnaround time for results and the performance to assess the risk benefit profile and to determine whether a device is substantially equivalent.

FDA acknowledges that point of care studies for troponin devices are difficult to execute. Point of care environments are busy and operators already have multiple tasks that need to be accomplished in addition to the investigational testing. This easily gets complicated by specific clinical trial requirements. For example to test different sample matrices or to match serial samples for standard of care with serial measurements for the investigational device. There are many instances in a point of care clinical trial where things can go wrong and as a result manufacturers may not get the data they need to support a pre-market submission.

The following slides describe some of the challenges that FDA has observed in these studies. Central laboratory troponin assays currently on the market are typically intended for use with plasma serum samples. In a point of care environment however use of whole blood is also desirable because it does not require much processing and is therefore easier and faster to use. However sample matrices may perform very differently in a point of care clinical trial because of a variety of factors and this should be a consideration for the design of the study.

FDA has observed multiple instances where the performance of whole blood is significantly different from plasma even when a matrix specific cut-off is used for analysis.

For point of care devices samples are typically measured immediately or very close to the time of collection. For logistical reasons clinical studies to assess point of care troponin devices may include samples that were stored for various amounts of time at different temperatures between collection and testing. As we discussed in our earlier session on

pre-analytical considerations sponsors typically submit sample stability studies for FDA review in order to bridge the performance of stored samples to the performance of fresh samples that intended users would obtain with the device.

FDA has observed cases where troponin is not stable in patient samples that are stored prior to testing with a candidate device. If samples are not stable then the reported clinical performance may not be valid.

Point of care environments are very different from central laboratory conditions and typically show greater variability and less control over operating conditions like temperature, humidity, et cetera. FDA has seen data in pre-market submissions suggesting that troponin devices can be sensitive to changes in the environment and that this can be a challenge especially when validating point of care devices. For example in a hypothetical pre-market submission the proposed labeling may warn of a significantly different result if the device is used outside of an operating environment of 20 to 24 centigrade. Such a narrow

range of temperature is difficult to achieve in point of care settings and consequently in our hypothetical example this led to the exclusion of 35% of samples in the clinical study.

Another consideration for point of care trials for troponin devices is the potential for differences in performance between clinical sites. FDA understands and expects that there will be differences between clinical sites. However, FDA has reviewed data where the performance at one or more clinical sites is considerably worse than at other clinical sites in the same study. In such cases it is not always clear whether the device is the cause of the poor performance or whether there is something about the site or the clinical study design that influenced test performance. For example poor performance could be due to differences in how the site handled, processed and stored the investigational samples. It could be due to differences in the patient population which may show that the assay's cut-off may not have been established well enough to overcome demographic differences.

FDA has also observed poor performance due to

the biased collection of different sample types. When there are many confounding factors in a study and performance issues arise this can greatly complicate FDA's review since it makes it difficult to assess what the performance of the troponin device will be in the intended use population.

Some of the questions that we would like to discuss in this session are: What expectations you as the stakeholders have for the performance of point of care troponin devices? What, if anything, should manufacturers include in their labeling to aid users at moderately complex point of care troponin sites? What are some of the challenges encountered while planning or executing a clinical trial for a troponin device and how could those challenges be addressed? And how limited is too limited for point of care troponin devices that are intended for use in just a narrow range of operating conditions?

At this time I would like to introduce Kellie Kelm, our Deputy Director who will moderate the panel discussion. And please ask all panel members to introduce themselves.

Thank you.

CLINICAL TRIALS FOR POINT OF CARE DEVICES

DR. CHANG: Hi, there. I'm Anna Marie Chang. I'm in the Department of Emergency Medicine and I run our clinical trials oh and I'm at Thomas Jefferson University.

DR. JAFFE: Al Jaffe from the Mayo Clinic.

DR. McCORD: Jim McCord, Cardiology, Henry Ford Hospital.

DR. PEACOCK: Frank Peacock, Emergency Medicine, Baylor College of Medicine, Houston, Texas.

DR. SAN GEORGE: Hi, I'm Rick San George, Head of Clinical Affairs for the Abbott Rapid Diagnostics Division formerly at Alere, San Diego.

DR. CAPOSINO: Paula Caposino with the FDA.

DR. LESSARD: And Juliane Lessard from the FDA.

DR. KELM: Okay. Well, first, Juliane thank you for that great intro into the subject. So the first question I think we have is a good starting point. So for our panel what expectations do you have for the performance of point of care troponin devices?

Anybody, we don't have to go in order. Any takers? Frank, Dr. Peacock.

DR. PEACOCK: So there is Rob Christenson and Rob and I used to sit and talk at these AACC meetings for days and days about how they should be identical. And I thought well, that sounds pretty good and then I drove here in a big fancy car instead of a pickup truck and I realized does everybody really have to have the same car and the answer is no. And so I've come to the reality is we don't all have to have the same assays. I have a different job than Jim McCord does. He's really interested on the rule in being right and not screwing that up. So he wants relatively tight performance around the 99th percentile. Me I want rule out, that's at the far end of the spectrum. That's the low end of the spectrum. So if you can tell me that this patient has and I don't even need a number honestly if you just tell me the test is positive and they are low risk they can go home; that is all I need. So it is a different job and we shouldn't drive the same car.

DR. JAFFE: I take a different point of view.

[LAUGHTER.]

DR. PEACOCK: I was waiting for you.

DR. JAFFE: Not infrequently with Frank. I think the reality is that often point of care devices are being used alone at institutions that have to deal with both rule in and rule out in rural areas where they don’t have the ability to have larger labs. It's in those circumstances they need a broader portfolio. Now if you wanted to develop a specific point of care test to do niche applications then I'm all in favor you can have a Yugo.

[LAUGHTER.]

DR. JAFFE: On the other hand

DR. PEACOCK: Spoken like a cardiologist.

DR. JAFFE: If you want something that can do both works you need to have them be more compatible and in the ideal sense I think if we don't set expectations for point of care troponin devices at a high, very high bar they'll never improve to reach the criteria we want. I'll share with you that back in 1999 when the European Society and the ACC got together to talk about how we were going to use troponin we put some stakes in

the ground. We said a ten percent CV at the 99th percentile. It is just now with high sensitivity assays that we are finally reaching that. But we put that stake in the ground because we knew it was doable and because we wanted to push the field to reach that performance. So I think we need to have a high bar and demand a lot for what we want with point of care because often it is going to be used as a solitary device.

DR. PEACOCK: So you are talking to the FDA or the IFCC?

DR. JAFFE: Yes.

[LAUGHTER.]

DR. PEACOCK: Then we've got to take about 15 assays off the market because they don't make your standard.

DR. JAFFE: Well, they made the standard at one point in time. They now -- we are now talking about a new iteration of assays coming through and I'd say it this way, I would not advocate that the assays we previously approved would get approved again if they came through again.

DR. PEACOCK: I don't argue with you but what I don’t want to be is to continue being Nigeria and we've got to have assays that are better than today available but they don't have to be perfect.

DR. JAFFE: I'm not disagreeing with that either. That said the point of care assays need to improve. We need to push them to set the bar as high as possible and recognize that most of the time they are going to be used for both rule in and rule out at most hospitals.

DR. SAN GEORGE: So we all want to have lab performance at the point of care so I think that is a given but we have to recognize that point of care does give one benefit over the lab and that is in turnaround time. And so I'm interested to know what, if any, tradeoffs might be acceptable to the clinicians on the panel or anywhere else where for that faster turnaround time some other level of performance might be an acceptable compromise, this is that benefit risk balance that we are all trying to find. We've talked about lesser precision in the past maybe that's not quite acceptable, maybe lesser analytical sensitivity,

does the measurement range have to be as high or as wide as the lab system. You know let's remember lab systems are based on analyzers that are hundreds of thousands of dollars, they do hundreds of assays, they can absorb those costs. We're trying to do the same kind of performance at a much lesser sort of system level. We don't even have a centrifuge so we do have to separate the blood without that as well. So given those challenges and given the quick turnaround time are there any compromises? Do we focus on ruling out those low risk patients and focus our performance in that area? Do we focus somewhere else? I'd like to understand what tradeoffs we might make.

DR. McCORD: I guess my clinical perspective on that would be is if you have to compromise and it sounds like you have to compromise it would be compromising more on the specificity and less on the sensitivity because this is used mostly upfront in the ED where you don't want to compromise on sensitivity if that's possible.

DR. JAFFE: And usually these sorts of devices are often used in smaller places where the level of

sophistication of the staff may be less and therefore the clinical component of this which often can save whether it is lack of specificity or sensitivity although I do think sensitivity would be more important in this instance is going to be limited and so you need to be very careful to at least be clear about what the metrics of any given assay are and what clinicians need to be wary about in the interest of patient safety.

DR. PEACOCK: We do that with D-dimer now. If you are negative and low risk you are out the door. If you are not negative then we have to do something else. And there's no reason that troponin couldn't be structured the same way. If you are low risk and only 30% of the people would fall into that category you are out the door. That leaves us 70%. We've got to do something else. Maybe we actually send another troponin to the lab to get a good troponin. But the idea is that if we have a rule out test and that paradigm does exist that would -- 30% doesn't sound like it is that many, it is three million Americans a year held in ERs now catching whatever they catch while they sit around and get coughed on by the guy next to

them. They could go home.

DR. JAFFE: I'm not sure that many of the present point of care tests are able to do that. There are some that are coming that have adequate sensitivity to rely on to send people home. But I'm not sure any of the ones on the market now have that.

DR. CHANG: I mean for myself what I would love is now we have a lot of places, my place has a physician in triage who orders a bunch of test, you know, so I'd love for the point of care to be able to say okay, I can sit on this patient a little bit longer, you know as a rule out test. I think that is much more helpful versus -- and then also to have a range where it is like okay, well, I do need to correlate this with the lab as a second draw. I think having that correlation is really important for me as well. So to know how to interpret the first one and then the one that I'll send to the lab. And I don't really know, we don't have point of care anymore at my new job, my last job we did and it was really hard and then when we would call the cardiologist they'd be like well, what is their lab troponin. I'm like ah, this is

just based on our point of care. So I think we do need some correlation to both.

DR. JAFFE: There are a couple of iterations of this maybe we want to talk about. One is that in Europe they are starting to use some screening point of care in the ambulance on the way in. The idea being that it allows those people potentially if they have a good story and they are usually mostly in the higher-risk patients if they have an elevation they can go past the ED and directly into the hospital, saving some ED time and allowing them now to test whether or not earlier intervention in non-STEMI may be helpful. So that is one potential use that could be considered. I think the assays need then to be validated, when we took them up in a helicopter they didn't do so well. When they get bounced around in the ambulances they may or may not do so well. So some of those validations are necessary. But that is one place.

There is a second issue as well. And it has to do with the definition of point of care assays which I think I'd like to throw open. There is a point of care assay which from my point of view is something I

can give Frank to put in his pocket and he can go around and get a drop of blood and put in it. And that is a point of care assay to me. But there are developing devices that are near patient assays that you could put a small machine in an ED and some people would say that's a point of care device as well. They are fundamentally different however and the metrics of some of them could be substantially different. So I think that is at least another element of this that ought to be considered.

DR. PEACOCK: The challenge for near patient testing is how much skill is involved in running the specimen because emergency departments can certainly do near patient testing. I just don't have a lab tech.

The other point I would like to make is there is really no hurry to get a positive troponin because when there's a positive troponin you get in a line, it goes over there and you are going to spend four hours, getting your room and you go upstairs and you get cathed maybe tomorrow in almost all of the cases. And so to hurry that process up doesn't really get me much and doesn't justify the cost. However to hurry up the

process of ruling out there is a huge advantage to that. So there are different ends of that spectrum.

DR. KELM: So one of the questions that we would have then is maybe you can weigh in on this is what percentage of false negatives would you accept with a rule out device?

DR. PEACOCK: Yeah, so like Dr. McCord said it is very -- this was published a number of years ago by Martin Than, he surveyed about 1,000 emergency docs and said well, what will you accept for an error rate? And it came out to be about something on the order of one percent or less. So you've got to have sensitivity of 99% and 100 would be cool but we'll never get that, but 99 is sort of the threshold for emergency docs.

DR. KELM: Okay.

DR. JAFFE: But let me push back actually in the other way because we have many cardiac troponin lab tests where we don't have that degree of error free performance and we rely on our clinical expertise to augment that. So what would it be acceptable to have a -- I think it would be helpful to have a point of care assay that provided the same clinical performance say

as the present-day laboratory assays and said, Frank, one of these days you are going to have to be a clinician.

DR. PEACOCK: And we have that data on the risk scores. We just finished a 30,000-patient study with EDACS and a pair of troponins and you can go home. And there's the heart score. So there are clinical metrics that already exist that make that a relatively easy step.

DR. KELM: All right. Great. So the next question what should manufacturers --

Dr. Christenson did you want to speak? Yes.

DR. CHRISTENSON: Just to ask the question. It seems to me that troponins run, if you are going to call the test troponin; right, that it has to be -- it can't get a pass at point of care for performance. But that is particularly true in the sensitivity side for what Frank is saying, what all of you are saying which is use it for rule out.

I guess what I wanted to ask and maybe at some point you guys could talk about it. Do you think that the high sensitivity assays, if you have high

sensitivity troponin that was at the point of care do you think that that would be an important advance because we're talking about shorter and shorter times between measurements. And if you send it to a lab and you have to wait for an hour and a half or two hours to get it, that delays the disposition of that patient versus a 20-minute turnaround time at the point of care?

DR. McCORD: And the short answer is yes, I think that ideally you'd like to have the high sensitivity point of care because these rule outs now are with high sensitivity at presentation one hour where the turnaround time now is a much higher percentage of the whole operational time encountering that patient. So that is why point of care is not going to go away. It is going to become just more relevant in my mind because of how quickly we have to assess people in ED now.

DR. JAFFE: Some of that may be helpful but on the other hand I think if you start looking at how fast the nurses can turn around a bed in the ED, how long it takes for the clinician to get in to really get a

history and be comfortable they know that patient to look at the ECG that some of this rush to make decision and rush to troponin is excessive and that it's actually more on the margin than a real effect many times.

DR. PEACOCK: Okay you asked for some politics because this is how this works in the ER is that about four hours our Press Ganey scores start to decline, that is the break point. And it has been studied in over a million patients that you go from being 96% high quality to 75 and so the hospital administrators are looking at you going you got four hours buddy. Because after four hours we get bad Press Ganey scores and then Obama Care sends their people down and our patient determinant margin shrinks by one percent. So patient satisfaction is now part of the game. And it is four hours.

DR. APPLE: So I want to follow up what Rob said is the concept, one comment, one question for you all. So do you realize that the large majority of point of care assays will test negative 15 to 20% of the time that you'll get a positive result in the

central lab. So that's the status right now of the world.

The second thing is how many of you have actually used point of care and -- okay. And then the reality of it is do you report numbers because I know your numbers are different than the central lab where they use T or I and how do you deal with the performance with your clinicians to give them a positive negative and part of that question is what percent of those point of care results are repeated in the central lab because they don't believe the positive or don't believe the negative.

DR. CHANG: I think that was the point that I was making earlier in that so my job right now doesn't have a point of care but my old one anytime that we called cardiology with a lab result or you know hey, this person has a negative they always wanted a lab draw anyway. And so what would happen is that you know patients would get some labs done point of care and still repeated and there's been studies to show this happens also with lactates and with every other lab that we do point of care there's always redundancy and

there's a big push I think by Arjun and some other people in emergency medicine to try and reduce those redundancies. But it is still not happening.

DR. JAFFE: And you can get into a long discussion as to how cost effective it is if you've got both up and running. I think the bigger use of point of care is that there are lots of place that don't have the ability to have a central laboratory, like it or not. And those places are totally dependent on point of care and on clinical triage. And that's why I think we have to be very careful because those patients are at equal risk as is patients that are coming into the Mayo clinic and yet those operations by and large have less sophistication and now you are giving them a much less sophisticated troponin assay. So I think we need to separate out the idea of what does Frank need to help him on X number of patients in an academic medical center from the larger number of people in rural areas who really rely on this on a constant basis.

DR. PEACOCK: So Fred, I'll tell you I live in the scenario you described. I have the Abbot i-STAT at the front door and we have the Siemens Vista upstairs.

The lab is upstairs at my hospital instead of the basement. But when patients come in I don't order the troponin; the nurse does the troponin on protocol. And so I get troponins on all sorts of crazy people that I really didn't want a troponin on. It's an 18-year-old girl who had 37 seconds of chest pain. It is like do I really need a troponin. And when it comes back negative I go your risk score is extremely low so we are done here and you can go home. But there are other people who are higher risk. They'll get two troponins. They might get them on the i-STAT but if anybody has a positive they end up sent to the lab because they don't trust it. So we get double troponins and there is a cost associated with that. But the objective is to increase the specificity. And then you have to use your clinical brain as well. So there are a large number of patients on the higher side who get multiple levels. And because the inside people meaning everybody who works upstairs, the cardiologist in the hospitals don't trust the troponin like Anna Marie said they want another one. So we do that.

DR. JAFFE: Frank, but doesn't that give you

trouble because you don't have a baseline to look at the delta?

DR. PEACOCK: Yeah, absolutely. But when you are 18-years-old with 37 seconds chest pain I don't really need it.

DR. JAFFE: 38 would do it for you?

DR. SAN GEORGE: So just a comment on the need for correlation if you will between the point of care system and the lab method, clearly that is desirable. You know and at the same time clearly a point of care method can't agree with all lab methods given the lack of standardization variability among them. So that is a challenge that we all face. The 99th percentile should mitigate that risk, that is you know if everybody used a 99th that correlated or was concordant with the others then some of that goes away. And of course a universal sample bank like the AACC or some other one of that type where everyone used the same sample set to establish the 99th might help in correlating that point of care method with the lab methods.

DR. JAFFE: But we're talking about point of

care assays as if they are still the old ones. And there are some newer ones coming and I do think that the newer ones whether they reach the criteria for high sensitivity or they are just short of the criteria for high sensitivity many are comparable to what central labs provide now. That's a different situation than the ones that we have, many of which are just terrible, terribly insensitive.

DR. KELM: So it has been a really interesting discussion. I like the idea of -- or the concept of for example needing point of care for those sites that need a sensitive and specific high performing point of care but then I think the idea of just to rule out a device that only does that, there are concepts there that would be interesting.

So I'm not sure whether or not there is anything else you want to add in terms of the second question. What should manufacturers include in their labeling to users at moderately complex, your point of care users or your sites that use point of care devices. I don't know if there is anything that you want to add.

No, I didn't think so. So for those of you who have planned or executed clinical trials or helped with that for point of care troponin devices have you encountered any challenges that you'd be willing to share and or thoughts on how you think you could address issues with trials?

DR. PEACOCK: I like the first line of our slide that said point of care trials are hard. I would argue with that. The key is who you get to do your trial. Every doctor will tell you that they will enroll 1,000 patients in two days. They are lying to you. They have a conflict of interest. They want to say yes. So if you ask their study nurse how many patients they will enroll, that is the truth because she's got no reason to lie to you. And if they don't have a study nurse that is not a site that works. So you need to have an experienced site with study nurses preferably there 16 hours a day. If you have a study nurse that's there 40 hours a week that becomes really hard to enroll because most patients with chest pain come in between 3:00 and midnight. If your nurse is there 9:00 to 5:00 and you want to do a zero, one, four

and eight-hour troponin you are not getting there. None of those patients get enrolled so your enrollment rates are terrible. There are about four ERs in the United States who do 24/7 enrollments. They are not very common but most of what I call the professional emergency medicine research sites and there's about 40 of them have something on the order of 16 hours of research nurse coverage a day. Those are where you go. And if you can access that this is not hard at all. The only hard part is those finger pokes ones which I personally don't see any reason why you want finger poke blood. But just because patients hate having all their fingers poked. But other than that they are easy.

DR. JAFFE: Well, there also is an issue of who does the testing. And the reality is it depends upon and it is site specific. It is not that any of this is hard. I probably could even be trained to do it although I won't push that concept. But it is not terribly difficult. On the other hand people who are very busy and who are running around and particularly if things take critical timing may not do it well. May

not attend to it well. If they're nurses they may get called away. So whether it is a nurse or a lab -- person who is laboratory trained or responsible to the laboratory has to be somebody who really understands how to do the testing properly. And I think often many sites have signed on when they don't have an understanding of the fact that you need to have availability, patients, a modicum of training and a modicum of dedication to how it gets done.

DR. SAN GEORGE: Yeah, so you know the big difference too is it is whole blood in many point of care cases and so the testing has to be done prospectively on site. So you need the right operators, you need enough of the operators, you need the intended use operators, you need to test multiple time points, you need around the clock research people if at all possible to do it right and to do it well and do it in a reasonable amount of time. So that is another challenge, if you will, is doing real time testing with whole blood. We talked a little bit earlier, it has been mentioned sample matrix effects. It would be nice if one could demonstrate a whole blood

to plasma equivalence and so that the testing could be done on plasma only in the study even on banked frozen plasmas maybe retrospective leftover samples. But in the absence of that being acceptable testing fresh whole blood is the only option and it presents some additional challenges.

DR. APPLE: Since I flew here I might as well speak. So Dr. Jaffe the definition of point of care testing is operator based, not the size of the instrument, even though we'd all like something in our pocket. If I'm correct that is the definition. So therefore my suggestion to the FDA is when at least this is what we're told by the manufacturers the FDA requires that to do the study they can't be laboratorians. They have to be done by true non-laboratorians because that's who the point of care is defined after. But one of the things that we've run into in our studies is we have nurses. We are 16 hour a day operation. But they don't allow, the manufacturers say the FDA doesn’t allow our nurses to run the point of care testing in a lab environment. So we have to find sometimes a closet or some out of the way place to

do it. So I would recommend who cares where the test is actually done. If there is a space in a laboratory, like we have a lab in our ER, if the nurse could walk into our lab and run the point of care that should be an acceptable part of being able to do the study trial. It shouldn't be limited to some place because sometimes the only places the nurses can do is step into -- we have a lab or space that could be laboratory oriented. So something to consider is that the nurse or the non-laboratorian that partakes in the study that does the testing doesn't have to do it -- can do it anywhere. Who cares what the site is?

DR. KELM: well, obviously the idea is to not always perform this in the best-case scenario but getting in the real-world scenario or worst-case scenario. And so we try to find out what from the manufacturers where they're intended to be used and who by and we ask them and generally we'd like to have the sites I mean to have various trained people so sometimes they are nurses or different staff and different kinds of locations. I mean we are always willing to talk but the issue with a laboratory is that

it is better controlled. Then the problem is the information on the label in terms of the performance going to be what Dr. Peacock would expect in his ED when his nurses have to go upstairs to the lab to test.

DR. APPLE: I'm not talking upstairs in the lab. I'm talking about a little -- sometimes we have stat labs but they might not do a certain point of care. So I'm just saying we have nurses that step into our lab -- we have a little box, our lab is about the size of this space, they'll walk into do a glucose. So I'm just saying is don't limit where the testing for the studies are done because it is the end user that's doing it, who cares where it is done. So you'll have a lot more opportunities, if the FDA has walked around and ER at all and see the space considerations and how busy, it is hard to find space sometimes to do these testings. They have to walk away from the exact site. Just a suggestion.

DR. KELM: No I think if you know you're talking about small spaces in ED for example small labs because yeah some of these also are larger instrument like little table top things that obviously aren't

going to be -- you are not going to yank it out of your pocket at the bedside. And you may have a little space in the ED where that is set up. I mean then that would be a valid place to do the testing.

DR. APPLE: Just keep an open mind because I think you could enhance enrollments and the processing a little more efficiently if you eased up that requirement that's all.

DR. KELM: Yeah. It’s not a requirement.

DR. APPLE: We're told it is a requirement by the FDA. When we do the studies manufacturers tell us the FDA requires the non-laboratorian to do the test outside of any lab space. So I'm just sharing with you.

DR. KELM: I think we define it as sort of a central lab location. But then, yes, if there is lab space where you do point of care testing in the emergency department then that would be a place of point of care testing is done in the emergency department.

DR. VERBARG: I'm Jasenka Verbarg. And this is a question for FDA requirements as well as I'd like

the panel's input on this. For point of care troponin tests specifically high sensitivity as far as having controls on the test say it is a cartridge that would take patient sample what is the requirement, what is the preference either having a control on the cartridge or a set of calibrators that would be used either on a separate cartridge that is specifically for control or on the same cartridge that we would test just from a batch?

DR. KELM: Go ahead.

DR. LIAS: So there are some CLIA requirements surrounding quality control testing. And some state and local requirements also across states that vary for quality control testing. So facilities need to fulfill a CLIA requirement with respect to quality control testing and external controls are typically available even for unitized tests. And often the recommendation relates to testing within lots or at some sort of frequency within the lab. You also have to pay attention to storage requirements. We also don't want to confuse point of care with waived testing. So here we specifically said moderate complexity point of care

testing because we didn't really want to address waived testing today. The question of waived testing is a little bit different because in that case the laboratory, the waived laboratory or the waived users would have to follow the instructions exactly which typically would specify more directly quality control use requirements. So it does depend on whether you are talking about waived or moderate complexity. Most of the requirements that relate to this are CLIA requirements.

DR. SAN GEORGE: Courtney just as a follow up then is there no expectation for onboard QC for a unitized device?

DR. LIAS: There's no requirement for onboard QC. A lot of times manufacturers will design some sort of quality metric just because it helps the design of their device. But external QC is separate from that.

DR. KELM: So last how limited is too limited for point of care devices that are intended for use in a narrow range of operating conditions?

DR. PEACOCK: Do you have an example?

DR. KELM: So I think the example is that we

gave in the talk for example was a device that only worked well between 20 and 24 degrees Celsius and indicated that outside of that range you could have greater than ten percent bias in results for example as a hypothetical. So you know we see this often with some of even not troponin but other devices like glucose meters, obviously then what is the operating range. If someone is using it on a soccer field or in sunlight and it doesn't work.

DR. PEACOCK: So I have to --

DR. CAPOSINO: So in the example the test result was not reported. You just got something like moved to a different temperature.

DR. PEACOCK: So I had to convert centigrade to Fahrenheit so I understand those degrees. And in Houston they air condition us to the point we are about to freeze to death because they have to prevent mold. So it has to work within the range of humans of an ER because this is where I think most of these are working. Allen said they don't work in the helicopter and I realize they have trouble in ambulances too for wiggling and ambulances can be 120 in the summer and

minus 20 in the winter. So it is an unbelievable temperature range.

DR. KELM: Well, I think the ambulance would be waived; right? Yeah, we do ask people who are interested in waived devices to try break those actually.

DR. JAFFE: But I think they need to be tested where they are going to be used. So if indeed you are going to use it on the soccer field and in the emergency room then there are different criteria than it is just going to be in the emergency room in Houston, Texas. If you are going to use it in the emergency room in other places that have saner temperatures you know not Rochester, Minnesota, for damn sure, then you may have a different temperature range. So I think it is the intended use can define that. So that a company could decide although I don't think it would be wise to have something within a very narrow range but then they'd have a very narrow intended use because they'd have to define the number of places that would work. So I think you have some flexibility. I think the only problem is that it is not

so clear once it's something that's approved that you can -- you really know where it is being used and could actually in some way implement a regulatory stance about it.

DR. KELM: Well, I think our thoughts are 20 to 24 might work well in the spring at some sites but then you purchased but what happens if the conditions obviously are here in Baltimore very well air conditioned in the summer and not so much in the winter, et cetera, et cetera. Maybe difficult for labs to control if it works well one season, not in the other. Seasonal variability that is going to wind up --

DR. JAFFE: But I think that is something that if a company acknowledges it only works from 20 to 24 degrees then the intended use is only going to be in those situations where it is 20 to 24 degrees and that may not be Baltimore in the winter. Now that would be a silly position I think for someone to take but I think you have much more flexibility based on the intended use. The real problem is how you really enforce that going forward once something is approved.

DR. McCORD: The practical issue is that some

old hospitals when the seasons change and our hospital is over 100 years old as the temperatures change it takes a while for the part -- areas to actually get warm or cold, so there can be quite a lot of variability. But I guess I have a question back to you because I sort of heard that the device itself may be able to say it is not functioning, it is out of its range and so if the devices can tell you they are out of operational range or they say you are moving too much or whatever it is then how much of an issue is it if it can communicate back to you that it's in a condition itself that's out of it functional range.

DR. KELM: Well, I guess our question to you is sometimes people might argue just label it and leave that. But I think the question is for those of you who use these in emergency department is that really sufficient or would you know about that? What would happen if it didn't work?

DR. CHANG: I mean I can tell you my nurses don't look at any of those labels. So unless the device told me that it's not working and it's out of range they'll still run it and they'll tell me oh, the

troponin is five and it's because the temperature is 32 degrees now in my emergency department because we just had a brown out. So I don't think unless the device can report back to us that it's out of its operating conditions then it is not going to work. And I don't know -- it doesn't make sense for a device that is going to be in the emergency department to have such limited capabilities. And things like humidity that was a big deal. We did a trial with a company and the humidity, my research coordinators were opening and closing doors and like dumping out buckets of water every two hours. It was pretty bad. So I think having a device that can withstand all of these conditions is very important.

DR. PEACOCK: So I would set the market -- the market should make those decisions. You have a device. It has to be in this range and I agree with Anna Marie that nobody reads this -- the nurses don't, if they drop it, they lick it, they stick it back in the machine. It gets really rough use. But if the machine says it's outside of the range then it is outside the range and that make it very easy to be compliant. And

if I want to buy an assay that only works for 20 to 24 degrees, let the buyer beware. That then becomes my problem to keep my lab that temperature.

DR. KELM: Juliane?

DR. LESSARD: So I think we also brought that up as a challenge for our clinical trial itself so if you have a device and it doesn't give you a result if you are outside the temperature and it is such a narrow range there's going to be a lot of error messages in that trial and then there's going to be a lot of missing data, there has to be retesting going on. So what do you think about that? It is one thing -- I think it is two separate issues one it is the trial and then it is the use of the device later on.

DR. PEACOCK: So going into the trial you know the device has a 20 to 24-degree range? Then they have to do it in 20 to 24 degrees otherwise you get errors. I mean you just go through the trial. I mean you wouldn't want to do the trial outside the range because you've already established it doesn't work.

DR. LESSARD: Right. Right. But if it is difficult to maintain that kind of range at the point

of care site like in our example then you would run into trouble with missing data because you are getting error messages all the time.

DR. CHANG: I mean as a trial site I'm happy to run anything; right. I mean when a company says please do this, we will do it and we will you know try and keep it within that range. But at the same time we do give feedback also to the company to say that's probably not very feasible. I've -- the last time we did this and I said you know, yeah, we're dumping out water and we're clearing our dehumidifier every two hours. I think the company did get that message but yes, if a company says these -- you know we'll do it. I am -- another of the sites like Frank said that has 17 hours a day of coverage so we make it work but it is not ideal.

DR. SAN GEORGE: So what if the device had a temperature range of 20 to 24 or whatever it is and if it were outside of that range, it told you it was either high out of range, low out of range and but still give you a result which you could interpret because you know if it’s high out of range if the

temperature is high the values tend to be higher or lower whatever the labeling would say. Same way with humidity. So you don't get a lot of errors or at least you don't get a lot of error messages but you get maybe a higher percentage of results that are obtained out of the ideal range.

DR. McCORD: I would say yes for the reason Frank said, you wouldn't want that. I think because there is going to -- he is going to get that number, he is not going to know it is out of the range. He wants it to be just not readable, this is not a --

DR. KELM: We've had some experience with devices with other analytes where staff actually doesn't know how to interpret those codes. And so we have had recalls where they will unfortunately lead to injuries or other things because nurses won't know and they'll just use it and not know what some of the error codes mean and they'll just move forward.

DR. PEACOCK: So for the point of a study if I have to get 500 data points and every third data point is out of range, well, I'm just going to get 1500 data points and eventually we'll get enough and I don’t care

about that because I'll do the research. And then when it comes out as a device you sell it and whenever it is out of range it gives you a nothing, it says unreportable, and you start to hate the device and you go buy the competitor. But I don't know that this is really a regulatory problem. It is just a research problem that you just have to get more so you can get the numbers and when it is done you can't report it.

DR. McCORD: It is regulatory in the sense that probably agree that if it shouldn't give you a number when it's out of is operational range. It seems like that is sort of inappropriate.

DR. CAPOSINO: And I think that is what we are hearing that you wouldn't want to report something that you have to interpret in a way following some labeling that may not be looked at.

DR. PEACOCK: Doctors trust numbers and if they are wrong we don't care we trust them anyway.

[LAUGHTER.]

DR. PEACOCK: So there has to be a trust there.

DR. JAFFE: Well, but the assumption sort of is that we'll all adapt to all these little things and

the truth is that many places whether emergency rooms or hospitals don't have the band width to appreciate any of this and they'll just go with whatever it is they have. So I think there's a real risk that people will ignore these sorts of things and even these sorts of messages. If the machine doesn't work at least then you say well I can throw it away and get a new device. But if you give a value I think it is often that clinicians will ignore the fact that there is an analytic problem.

DR. ENGINEER: I would take a patient safety perspective on that question that Rick brought up that if this was my mother and she came in and someone misinterpreted, what is the change of someone misinterpreting a value that was out of range but reported. I think there is a high risk of having clinical error there. So I would definitely avoid that and take the whole discussion from a patient perspective as opposed to a laboratorian, cardiology or emergency physician. How would we as patients want to take that and would that how we would want to be treated. So I think that a lot of those things bring up

some concerns there.

DR. KELM: I think if a lab, you are talking about some of the rural locations that if this is their only device that they are using and if a lot of the time it’s not going to work at the first or second or third time you run it then what happens in those locations don't have a backup method or a clinical lab method. Hopefully they don't buy it. But --

DR. PEACOCK: But having a bad lab is no different than having no lab.

I mean it if it says it’s too warm then you cool off the machine. And but getting the wrong number I - doctors will do, they look at the number they'll just do it. They'll say oh, it's positive because it says it is positive. The fact that they know it is out of range they won't even think about it.

DR. KELM: So I guess I'm freezing if you actually a device in the temperature range where it doesn't work and it gives you just error codes but if the problem is that these machines will and you are in a location where it is the only thing that you have I mean obviously that winds up being a patient safety

issue because you actually don't have a device on hand to use at that location. And if they don't read the inserts and don't read about the device and it doesn't know that it doesn't work that way before they buy it then --

DR. JAFFE: But I would think that part of the labeling of such a device should be very clear with whether it is a big black stamp or a big red stamp or some other color that makes it very clear to whatever user and I think that is something you could do easily with labeling.

DR. WIENEKE: Hi, I wanted to bring up a separate challenge that I've come across if we have a few minutes. Do we have enough time Kellie? So one of the challenges that I've come across when looking at some of these clinical studies regarding point of care is the clinical sites in which the devices will be used. And we sort of touched upon it. But I'd be interested to hear from the clinicians what your anticipation is as to where these point of care devices are used or could be used and what type of site should be used to determine the clinical performance of the

device. As an example so Dr. Peacock brought up the fact that he can do these point of care studies in his emergency room that has 16 hours shifts. And I'm sure point of care devices are used in the emergency room. What other types of sites are these devices either used in today or going to be use in? And do we need to require the clinical studies to be performed in those sites. For example some of these free-standing emergency room places that pop up. I'm assuming some of those -- now I'm not familiar enough to know if you show up with chest pain do they automatically put you in an ambulance and send you to a real emergency room or are they going to do a troponin, evaluate you and do a troponin test. I'm assuming either a handheld or benchtop because they do not have central lab. Do we need to have the clinical trials include sites like that? Or could you imagine such a handheld device being used in a physician's office where because once these things are cleared we have no control as to where they go. Could a physician's office have one of these handheld devices and if a patient shows up with chest pain want to do a troponin, is that a crazy idea? Or

is that possible? And do we need to consider those possibilities in the clinical trial designs when we are getting the submissions. And I just raise it as a question because these are the questions that I have when I look at the clinical trials. If it is always done in emergency rooms and I know Dr. Peacock is doing it okay it's a reasonable study. But do we need to get out to some of these other sites that possibly the devices could be used in?

DR. JAFFE: See I'd argue that in the sophisticated emergency room Frank needs to bug his lab and look at his processes and make sure that his turn around time is a little bit better so he can get rid of that. In terms of the real utility of this it seems to me that there are places that really rely on that and those are the sites where it ought to be tested. If you are going to have this as the only device in a rural area then you need some sites that really recapitulate the difficulties that that rural area is going to have. Most of us would not recommend at present and maybe with high sensitivity eventually it will change physicians having instruments point of care

or otherwise in their offices. All right. That doesn't mean we always listen. But and I think most of the small boxes do transfer people who they take seriously. But I think that are in some big rural area and that's where the studies ought to be done because that's where I think at least these sorts of devices ought to have some function until we get good enough to have the really good assays at the point of care that don't take that sort of care and feeding.

DR. CHANG: I would also add urgent care centers would be a big one for me because Judd Hollinder has all of us as emergency physicians also working at our urgent care centers and we have now, it used to be a lot of young people with sprains, strains and everything else and now I'm getting the 80-year-olds with chest pain or you know the 65-year-olds where it is like it is probably not but it would be great to have a device with that low detection limit that I can use and say otherwise they are also getting transferred to the emergency department. They're getting two bills. So I mean for patient comfort and experience issue I would love to be able to get a point of care

device that says, okay, it's not, this is really just your muscular/skeletal pain and be able to send them out.

DR. KELM: I'm sorry. Is the urgent care attached to the hospital or --

DR. CHANG: No. It is separate.

DR. KELM: But do they tend to then be transferred to hospitals?

DR. CHANG: So I mean I've worked in the past week and we've had multiple patients transferred because we don't have a point of care troponin device. And they are like -- and literally the note from our doc is needs a troponin and then sent home.

DR. McCORD: You can do EKGs at urgent care; right?

DR. CHANG: We can do EKGs. We do have an i-STAT now but no kind of further testing. So like for the physicians who also want a D-dimer and stuff they are transferring them.

DR. McCORD: It's a risk sitting next to Dr. Jaffe and disagreeing with him a little bit. I would say that there is probably utility for this in big busy

emergency departments were that turnaround time can actually have a practical clinical impact so I would think these trials should be done in some big hospitals and in a mix with the small hospitals where they are used. I mean wherever they are going to be used you should try to have some trial with that. I've heard them being used on cruise ships and aircraft carriers. And I don't know if you can do trials on aircraft carriers but wherever they are used it seems like that is where you would want to study them.

DR. PEACOCK: So what you're looking -- to give you an answer to this is what is the definition of someplace who might use it? If you can do cardiovascular monitoring in other words you have an EKG monitor then you could use serial troponins and have a clue what is going on. A doctor's office usually doesn't have that capability and I would not support these being used in a doctor's office because the danger here is you have a little bit of a heart attack, your troponin goes up, and then you go into V-tach and die and it can be over ten seconds. They just drop dead. And so if you can't handle that you should

not be testing for troponin. If you can handle that then, so you have a crash cart, a cardiac monitor, then if you want to do troponin testing then that's fair but you've got to also be able to hold them the length of time it takes and current American College of Cardiology Guidelines are six hours. So if you close in four hours that patient is getting transferred anyway why are you bothering testing.

DR. McCORD: No clinic is going to want to do that I think.

DR. PEACOCK: Huh?

DR. McCORD: No clinic. I think.

DR. PEACOCK: Clinic is fine.

DR. McCORD: I can't see any internist, primary care doctor wanting to do it. And if they do then you probably don't want to see that doctor.

DR. PEACOCK: But we have lots of free standing ERs popping up, Texas is sort of the epicenter of this. We have well over 1200 of them now. They tend to be focused around large cities so it is not a rural thing and they are truly free-standing emergency rooms that there's all sorts of political issues about

them scraping off the insured population and letting the uninsured go someplace else. That is a different animal. They are here to stay. They provide convenience to patients. And they are fully functioning, so I think they should -- it would be reasonable to test point of care in that environment.

DR. SAN GEORGE: So many of those environments I think and correct me if I'm wrong guys, rural hospitals, urgent care centers are not set up really to do research studies. So to the extent that that is true would it be acceptable for us as a manufacturer or sponsor to hire people, people who can do the consenting, people who can do the testing. Their intended use operators or selected to be representative of intended use operators but they can be put in those setting where we want the device to be tested on the populations where you want them tested? But they are just not set up to do research studies. Any comments on that?

DR. CHANG: We've done flu testing point of care device testing in our urgent care centers.

DR. PEACOCK: But to Rick's point it is really

difficult to do studies in non-academic environments. And academic environments tend to be large hospitals in large cities. And so when you want to have urgent cares and you want to have clinics and you want to have non-academic hospitals it is impossible because the reality is they see one chest pain patient every other day and you are going to hire a nurse to do research in that ER. It is going to cost you a ton of money.

DR. KELM: So we are up against the end of the session. And I do want to get Dr. Apple's comment because he's been standing there for quite some time. And I think that is all the time we are going to have.

DR. APPLE: I'll be real quick so Chris can speak.

DR. KELM: Sorry.

DR. APPLE: So I'm a laboratorian, none of you are and I think we're not -- putting aside rural practice which I support you have to realize a lot of your colleagues, your experts come to us and everyone expects the point of care and what I heard here is that Frank said it if you can't support it, you can't do it. But we just can't put point of care just because they

think it is an improved turnaround time because as you said from Jefferson no one believes the results and you repeat them anyway. So I'm advocating to the dismay of my industry colleagues until we have point of care in a either high sensitivity or at least equal contemporary if you get a one hour turnaround time I think your clinicians can live with it and why put point of care in there at all if you don't need it.

DR. KELM: All right. We will let one more comment.

DR. deFILIPPI: Thank you. I feel special. But I was going to suggest manufacturers working in evolving health systems, health systems are not hospitals. Health systems are systems they include outpatient setting, they include EDs, they include a lot of urgent care. So you can approach a health system, ours, University of Maryland, many systems and say I want to test in these multiple environments. I think that’s very possible to do that under the auspices of a single investigator and to reach out and do that. It is more capable today than it was five or ten years ago.

DR. CAPOSINO: I think I would just like to say that we want our studies to reasonably reflect where they will be used. We are not feeling like we have all of these restrictions exactly for you to give us a photograph of. These should reasonable reflect the intended use and the intended use operators. I think it is important to pick sites and to have the staff in place that you can get the information that you need if that makes sense.

DR. KELM: Well, I want to thank everybody. This was a very interesting panel. Thank you very much. So I think we are going to take a 15-minute break and we will be back at 2:25 for the next session.

Thank you.

BREAK

DR. CAPOSINO: All right. If we could start the fifth session. That would be terrific. If I could get the panelists to come up.

And just one more time if anybody wants to speak please register. I didn't look during the break to see if we've gotten more people registered to speak in the open panel. That would be great.

So during this session the panel will discuss the Use of Existing Data to Support Claims.

In this slide we just want to highlight our center's initiatives to promote the use of existing data such as real-world evidence and this is our guidance document that you may want to reference.

Our regulations actually allow for a wide variety of evidence to support the clinical use of devices. We're interested in good data that are available and we're also interested in efficient ways to support devices.

From our perspective good data are good data whether they already exist or whether you are working to collect it.

The discussion topics that we've identified for this panel are to discuss best practices for using existing clinical data. Discuss additional useful clinical uses for troponin, perhaps some that are not on our radar. And what sources of data may be useful to support troponin assays.

I would like to open the discussion panel and ask the panelists to introduce themselves.

Ian Pilcher will be moderating this panel.

Thank you very much.

USE OF EXISTING CLINICAL DATA TO SUPPORT CLAIMS

DR. PILCHER: I'm Ian Pilcher from the Division of Chemistry and Toxicology Devices. And before I allow everyone to introduce themselves I do just have one quick announcement.

Somebody lost a cell phone so if anybody does find one laying around please just take it to the desk out front. Thank you.

DR. CHANG: Again I'm Anna Marie Chang, Thomas Jefferson University.

DR. ENGINEER: My name is Rakesh Engineer, Cleveland Clinic Hospital. I'm a practicing emergency physician and just for a little background we've rolled out high sensitivity troponin at as of today seven sites within our 19-hospital system and we have a large database and that's why I'm here.

DR. GUTIERREZ: My name is Alberto Gutierrez. I'm recently retired from the FDA. I was at the FDA for 25 years, 17 in in-vitro diagnostics and I was the Office Director for eight years.

DR. RICHARDS: Hi, my name is Karin Richards and I work for a company called Precision for Medicine. And I have the benefit of helping industry interact with FDA on several regulatory and clinical related matters.

DR. SANDOVAL: Yader Sandoval, interventional cardiologist and assistant professor at Mayo Clinic.

DR. KELM: Kellie Kelm, Deputy Director, DCTD.

DR. LIAS: I'm Courtney Lias, FDA.

DR. PILCHER: So I think to start this you know we've had a lot of discussion about clinical trials. What should be in clinical trials, indications for use and rule in, rule out and different uses of troponin including the difficulty of conducting and designing these clinical trials. We've also spent some time on the analytical data.

But I think we would like to have an open discussion and get everyone's input on additional data that's already out there that is kind of maybe a little bit outside the scope of what we've discussed earlier. So I'm just going to start heading down the line here and see if you guys have any input on any additional

sources of data that you think would be useful for FDA both on how these devices are currently used and additional uses for troponin that we haven't seen yet.

DR. CHANG: So I think now with kind of electronic health records it is really or could be much easier to data mine. I think it would be easy at places like Rakesh's shop to get some large data sets and see how people are actually using these.

I mean additional useful clinical uses I think that there's been lots of papers out there in different disease entities and how troponin can help prognosticate these patients. So I think that would be especially these additional clinical uses like subarachnoid hemorrhages and things like that where it is not the primary indicator that may be a good place for some of these data to be used and started to kind of investigate these clinical conditions. I'll let Rakesh take it since he actually has high sensitivity troponins at his shop.

DR. ENGINEER: So I guess the place that we are using our data and my data comes from a different place than most other people. So most people structure

a trial, have an a priori question, get funding and have a research nurse collect everything. That is not what we're doing. We rolled this out in an effort to try to get patients treated in the right location and the right level of care. Those who needed to be at home should be at home. And that was the goal of our project. And all the data that we collect is in support of that project trying to determine what is the best protocol? Are we doing the right thing for our patients? Are we causing harm through changes in this protocol? And how does this overall look? So what that means though is that my data is somewhat limited. I have 1,700 patients who came in with chest discomfort or some other atypical chest discomfort type symptom. And from that a certain number of those received high sensitivity troponins. And based on that several of them were listed as low risk. From there another percentage of those patients were discharged to home and then another percentage were kept overnight. And so our database doesn't have the traditional demographics and doesn’t have the same type of follow up quality that we usually do. But we do have is

billing follow up and we know that we are not seeing a large number of negative outcomes when we look at our billing data. But it is outside of the scope of our normal data collection that you would expect. I am not sure if I answered the question adequately and if anybody has any other questions on that.

DR. RICHARDS: Okay. I'll go. Just to address there are kind of two pieces to that question. I think the first part about what other existing clinical data I think one of the tools that is really valuable and important to us is in industry is having access to cohorts. And there's a lot of great data out there and a lot of great stored samples that manufacturers would like to be able to use and we've discussed openly with FDA and they're receptive to the use of these leftover specimens.

I think some of the key things to really think about though in the feedback that we've received is really understanding that that cohort was collected in the intended use environment that you're seeking. So you are not just getting samples from a bank that is for a different intended use or outside the scope of

your intended use population. I think the other aspect is having to manage and make sure that you understand if you are missing samples from that cohort that you don't have biases in the samples that are missing versus the original cohort. And then of course sample stability issues. But I think the key at being able to demonstrate that the biomarker that you are seeking is stable in that frozen sample.

I think we've talked a lot about how expensive and challenging it can be to conduct clinical trials prospectively and so I think it is very important for all of us to have the option to go to some of these centers that have done these collections and be able to leverage the frozen samples. And obviously that doesn't work so well for whole blood but assuming we're talking plasma and troponin. And be able to use those in trials. And re-demonstrate that with your product.

I think the other question about other clinical uses for troponin, there are so many. I mean there's heart failure prognosis, risk, looking at symptomatic CAD, looking at general screening populations. So there's publications that can be used

for that purpose. There's again testing that can be done based on cohorts. So there's quite a few sources of literature to support alternative intended uses and we'd like to be able to also leverage the use of those publications to support additional claims because at the end of the day you are still just detecting the biomarker in the sample.

DR. SANDOVAL: So I guess the statement discussed for using existing clinical data I guess in my mind it could be interpreted twofold. I am a bit confused here. So maybe you can clarify but on one side we can be talking about how to best use the exiting data available in electronic health records to maximize potential research opportunities. And by the same token I guess how I understood the question was how do you use the existing published peer reviewed data available to how we are going to implement high sensitivity assays in U.S. practice. So in that regard even though we are I know over the next hour we will discuss about other potential additional useful clinical uses. I did touch briefly on this earlier but I think it is critical to continue to emphasize is that

right now the main use for cardiac troponin is to detect myocardial injury and aid in the diagnosis of myocardial infarction. And when you write that and say that out lout it sounds like a simple statement but yet it is not. And when you look at the literature there is a lot of heterogeneity in a bunch of studies. We put together a number of tables about the peer reviewed research and it is quite variable. And I think I at least wanted to take the opportunity to emphasize that. So right now we say use troponin to evaluate patients with suspected ACS or MI, it is a concern. However when you look at the literature and really delve into that you realize that some studies are primarily inputted solely type-1 myocardial infarction. But I think Dr. Jaffe mentioned before and I completely agree the study should probably be both type-1 and type-2 myocardial infarction. By the same token as you delve into the literature you see that some other studies they are actually talking about intention to rule out acute coronary syndrome. And they include unstable angina in their endpoint. So there's a lot of variability. So even though we phrase use the test to

rule in and rule out acute myocardial infarction how the studies had some of the existing clinical peer review data it is out there, it is quite variable. Some in type-1, some in type-2, some it is merger, some it is acute coronary syndrome. And I think there is of course a panel of clinical trial designs earlier but it is uncertain to me that we fully address in it in detail and when we are addressing that system clinical data right now that is an area that I think there is some degree of confusion. At least my personal opinion they should be all any myocardial infarction just not type-1 even though many studies have focused solely on type-1 myocardial infarction.

In regard to the intended clinical use as I said should be intended to detect myocardial injury. And I think one of the probably good documents, the first author is over here, Dr. Newby, it is the ACCF 2012 document on troponin expert use because if you look at that document it really phrases very well the whole evidence for all the conditions in which we use troponin for pulmonary embolism, for heart failure, for myocarditis, for cancer. And I think it phrases well

that right now the main indication is to diagnose to rule in or rule out acute myocardial infarction and that probably outside of that the most robust data it's either for cardiac toxicity in cancer and there is I guess there is an FDA indication for prognosis in renal disease. Outside of that if you look at PubMed you probably can find paper for any clinical condition associated with troponin. Whether that is actionable I am uncertain.

DR. KELM: Well and so you bring up something that I think sponsors as well as we struggle with because we are extremely open to sponsors using existing data whether or not that is samples from a study that's been done or we don't get as much with the health records yet. But obviously often people are performing studies because they have a question that they want to answer and they are designing it for and there is a reason why they are looking at one small group. And it is not designed to validate a medical device. And so that is I think a difficult thing for sponsors as well as FDA to use a study if we want if for this intended use but we are using a study that was

only studied in this small part of that group. And so then that winds up being very difficult and in many cases you can't extrapolate it to the whole intended use. And so that is often some of the questions that we have when sponsors use and come in with an existing study is how is that actually -- unfortunately it is very limited or it has this limitation and we struggle with how to use that to support something greater.

Alberto?

MR. GUTIERREZ: Yeah. So let me -- I see here two different sets of issues and perhaps let me address both of them differently. And I don't mean to be the one to put a hamper on this. But I know the agency is really thinking about how to use real world data and yet that is an issue that is difficult particularly for in-vitro diagnostics. Particularly it is going to be difficult for troponin when you don't have an electronic record system that actually tells you not only what value it was but what cut-off it was, which assay was used. And since the assays are not harmonized then you have variances. So your real-world data is going to be a bit of a mess in most cases. And

how do you leverage that with all that noise that that is going to have into something that you can use is going to be difficult. Perhaps you can get places where they either use the same device with a large number of people or where you have some stability and then you can use that. But then to leverage that for somebody else is going to be difficult again because the values are not normalized and how do you bridge to that. It is a difficult issue that I think is going to take a while to work out.

And to sort of extend the same is the problem even if you are using some previous studies and you have stored samples that you can try to use part of the problem here is that the studies themselves or the intended use that you want to have may require a different assay design. And we talked a little bit about that in rule out so for example in rule out you may want to have an assay that has an LoD but in other places you may not. So if you have data that you have collected with a particular assay but the intended use that you want now is going to require different assay design than what you have how do you bridge that. That

is also a difficult problem.

So using data that exists is great, it really saves you some effort but it does have a lot of the devils are in the detail, it does have a lot of issues as to how you bridge and what can you gain from the data that is out there.

Now there are areas and I do want to point out that the agency has done this a lot so there are areas where the agency has given totally new intended uses to devices based on literature data or based on data so hemoglobin A1C became a diagnostic device for diabetes. And that was totally done based on clinical data from published data, not from the companies themselves. So it is possible. But I'm not sure in troponin there is the conditions now for that to happen easily.

DR. LIAS: So I agree with some of what Alberto said because he emphasized that at FDA we really want to be able to leverage data. We want to be able to use data that are available that can support new uses. And we also want to be able to encourage the development of new pathways to generate and use real world evidence. So one of the things that he mentioned

is that currently we don't have perhaps an infrastructure to leverage some of the data that is collected using existing devices either in this country or in other countries. For example in some cardiovascular devices they have actual registries where every patient that uses that device certain parameters are required to be entered by the health care providers or the hospitals into those registries. And so there is data and those registries were actually designed to be useful and/or they've ended up being useful. To be able to find out how the devices are used, to understand whether safety signals exist or don't exist, to support other uses, perhaps collect data on certain patient populations, et cetera. I'm not aware of any registries or databases for in vitro diagnostics that exist like that and Alberto said the electronic health record doesn't actually currently-- one there is a lot of variability in health records. But even if you used a certain type of health record those health records don't necessarily collect the information one might want to be able to understand the data that you are getting out of that.

So one of the other things that perhaps we could talk about during the session is what can we as community do to build infrastructure to access real world data and or build the infrastructure to more efficiently generate the data to support either the current MI related uses for troponin or other uses for troponin devices that may come in the future.

DR. SANDOVAL: I think there are some changes that have recently occurred that I think are optimistic in what things could happen but I yet think that there are a lot of ongoing challenges that will limit the validity of research. So let me explain myself. Right now again for our main indication for which is acute myocardial infarction a large number of studies rely and I think this was discussed before but rely essentially on a team of adjudicators that essentially at least a couple of people look at the case to say whether there is injury or infarction, which infarction subtype that it is and go from there. And if there is a disagreement they can go to an arbitrator. Of course there are variations but that's a nov (ph) review. When you talk about what other large registries there are

and I can think of at least a couple like ACSIS in Israel or SWEDEHEART in Sweden, they have imbedded for example the universal definition of myocardial infarction within their registries so then you see that they have publications that have thousands of patients in which they've already adjudicated upfront prospectively for type-1, type-2 and provide lots of data. I don't think that right now you can pull that with certainty from the EHR not only because it hasn't been uniformly coded but because even within research investigation well-read investigators we don't agree sometimes which in what constitutes myocardial infarction and its subtypes. I do think there is some degree of use of how we can use EHR so starting this October effective October there is now, there is a Medicare approved ICD code for type-2 myocardial infarction. And if we agree that the menus for troponin will be to assess for both type-1 and type-2 then at least there will be some sort of observational way to extract from EHR both types of myocardial infarction now that there is a code for type-2 myocardial infarction. I do however would be upfront that we --

there is not a uniform agreement in what that exactly is.

DR. LIAS: So the use of ICD codes is an interesting one. And it has come up before. There is some indication in the literature that the accuracy of the diagnosis gained from ICD codes can be lower than we might want, sometimes in the realm of 75 to 80% accurate which could impact how we could get performance estimates out of ICD codes. I don't know if there is a way to address that issue.

DR. SANDOVAL: Right. So historically we actually published with Fred and some of the UTROPIA data on this. So that is the reason I say from data, from existing clinical data prior to October that type-2 was not part of that. I would say it is quite messy. So most of the ICD code marker unfortunately the large majority represent type-1. But when you look at some of the adjudicated events some of them might be type-2. And when you look at a lot of the events that were not coded as myocardial infarction but rather just as where troponin increases many of them include type-2 myocardial infarction. So it is quite messy. Whether

that would change now with a new dedicated code for type-2 myocardial infarction is yet to see. So I'm not sure that that prior data could be used, whether it could be changed prospectively we'll have to see what happens.

DR. LIAS: Another challenge we see people running against is when they try to leverage data from existing trials, Karin already mentioned that you have to deal with sample stability issues making sure whatever you are measuring was either measured soon after collection or was stored in a way that keeps the analyte stable. But beyond that when clinical practice changes significantly, when the definition of MI changes significantly we struggle to understand how that might impact the way that you could interpret the clinical performance of a test using data that were -- or patients that were diagnosed back in the 1990s for example. So that is another thing that we run up against is how can we leverage necessarily some of these large trials that were done 20 years ago. Maybe this will become less of an issue. But maybe only if we can get some new large cohorts available for people to

do some of these studies on.

MS. RICHARDS: I think maybe another way to look at this too though is if we're pushing the URLs down and I'm talking about high sensitivity troponin right now, if the URLs are really being pushed down in sensitivity levels and we've talked earlier today about what is a normal. And if we are testing other biomarkers to say this is really a healthy patient. So those get pushed pretty far down the detectable spectrum. Then if you look at perhaps in the literature where let's say for heart failure patients tend to have troponins that are elevated. Do you really need to prove that in a study or because you know that your URL is so low and your normals are here, down on this end, and heart failure patients tend to have elevated troponins by what's been presented in the literature do you really have to test for that or can you use the model to support?

DR. LIAS: Right. Sure. I'll talk about theoretical sense and we haven't looked at this particular question but theoretically if you have a biomarker of some sort, we'll talk about troponin in a

minute, if you have a biomarker of some sort where there is a lot of evidence in the clinical literature that a certain patient population has everyone with heart failure in this case would have a troponin that is elevated ten times the upper limit or something like that and that there is a fair amount of consensus on that point, we don't typically make people redo that. We would rely on good studies and the literature assuming that that is sort of a well-accepted thing. And typically these things are often already incorporated into clinical practice before they reach us in that case. And if it were true that the assays could sort of accurately measure that we wouldn't necessarily have the companies redo a study to demonstrate that that's true.

The challenge with troponin is that in some cases on one assay the values they are just vastly different. So troponin is a unique one because the value you get from one assay is very different from the value you get from another assay. So you have the additional complicating factor of having to try to understand what were the assays used in the studies and

how might that relate to the assay that you have right now. And maybe that will get better as the assays you know if this separates clinically if analytically they can measure really low and the assays converge a little bit that might help. Of if in the future these assays are harmonized however right now they measure different epitopes, some places you can have the same sample and one assay might measure five and another assay might measure 25. So that's not necessarily as easily interpreted depending on the situation.

DR. SANDOVAL: I do think it's important you are just basing on the discussion of heart failure and I mean again do you need demonstrate that there is a use for this whole array of circumstances. So I think it is tricky. I think the reality is that just to go back to what I said at the beginning it's the de facto test to detect myocardial injury. But myocardial injury can be identified in a whole array of circumstances that we would need a whole session to discuss the number of conditions that cause myocardial injury. So it is the de facto test to detect myocardial injury which aids in the diagnosis of myocardial infarction

which is the primary intent that most insert packages have for myocardial infarction to aid in the diagnosis of myocardial infarction. Do we need to have a separate claim for detecting injury in each separate condition? I am uncertain a little bit about that but there are recommendations for prognosis from different conditions, so class heart failure if I require I think it has a least a class one or two recommendations for prognosis. Whether many of these conditions for myocardial injury are actionable that's a different discussion.

DR. KELM: Well, there is still off label use and obviously if clinicians are comfortable with the assay they have on hand and using it for something broader than that. But obviously I think most sponsors at this point have been comfortable with a better controlled trial design with MI where that can be defined in a way. But I mean if people are interested in claims for injury or something else that they'd like to put that their device can be measured and used for we are happy to talk about those proposals and the data that support that if that is something they want to put

in their label as a claim for their assay. And then obviously we're open to whatever data that would be.

DR. SANDOVAL: And I would argue that that should be de the facto claim, that is what the test does. If the value is above the 99th percentile it is intended to detect myocardial injury, whether it is acute or chronic, or whether it is ultimately due to myocardial infarction is a separate clinical question and it will be used to aid in the diagnosis of myocardial infarction. But all troponin assays if they increase above the 99th percentile that is what they do.

DR. LIAS: So I would like to invite some of the manufacturers if they are willing to get up and talk about what kind of things would you need or would you like to have in an optimal world and Karin certainly weigh in please but also if there are other people in the audience from industry or PIs what types of real world evidence if you had an ideal future would you like to have to be able to either make trials easier, to make device development easier, to do something different that you don't have now. Because

one of the goals would be to figure out are there things that we can work toward as a community to make sort of innovation easier and to make these devices improvements quicker.

DR. deFILIPPI: So I'll ask as a PI. So years ago we measured N-terminal pro-BNP in a very well characterized NIH cohort. So the NIH longitudinal cohorts are meticulously designed, meticulously adjudged, outcomes often associated imaging. And we had done the study, the sponsor of the study thought great we'll take the assay and we'll use it for prognosis in older adults. And it came back from the FDA that there was -- the list was so long I think everyone was very dejected. Does it sound like let's say you take one of the newer cohorts from the NIH, again meticulously collected, associated imagining, adjudicated outcomes and you say we want to take troponin, take the general population, diverse population and say we can predict who is going to develop heart failure. Is that something that now is easier to come back and consider?

DR. LIAS: Well, not being familiar with what

the issues were before with the study. So you are proposing that there might be cohorts out there that somebody could use to say a baseline troponin could predict future development of heart failure. That -- I mean --

DR. deFILIPPI: Yeah, the NIH in particular has developed lots of cohorts for these at-risk populations. It could be patients with renal disease, the general population, heart failure and they couldn't be better designed.

DR. LIAS: There is no inherent problem with doing that. So if that is something that works; I think there are a couple of considerations here. When you are using existing data sets one thing you have to be careful of is if you need to set cut-offs or if you need to do some device development or design and I think most of the manufacturers are aware of this you should not go into the study that you plan to use to try to do that. So companies may also need data sets with which to design their device, the don't use up all of the data that they have because they can't reuse data they've used to set cut-offs to validate those

devices. So if you've designed the device separately and you want to use these cohorts to show how it works and the cohort is in the intended use population it sounds like that's the perfect type of study to use. So we are always happy to talk about those types of claims especially if people think it would be useful.

DR. PILCHER: Another question.

DR. BATES: Yeah, could you comment on the thought of having to finish the reference range study before starting a pivotal study in terms of thinking about how we might be able to do it a little bit quicker.

DR. LIAS: Technically you don't have to. You would need to prespecify that that is what you are going to do and you wouldn't want to look at your data ahead of analyzing but technically you do not have to finish the reference range study before doing your clinical study. I think what you do have to do or should do before doing your pivotal study is those analytical performance evaluations we were mentioning earlier knowing the analytical performance parameters that are going to be critical for doing your pivotal

study like sample stability and sample handling pre-analytical steps; that part is necessary to do before doing your clinical study.

DR. PATRU: Hi, I'm Maria Patru from Ortho Scientific Affairs. I have a question regarding the additional claims. So we as manufacturers we would like to perform -- I mean to have a lot of claims. The ones that make sense in the clinical practice, however, it is not practical and we also do not have the expertise to conduct such studies. So we shoot usually for a main claim, for example, MI right aiding the diagnosis of MI for troponin and we would like to have additional claims that make sense clinically. However, as I said we do not have the expertise, nor do we have the power, financially and the time to conduct I don’t know four studies to launch an assay. So my question is from the regulatory perspective if there is a study or multiple studies out there that experts, clinical experts recommend to the FDA to be considered is it reasonable for the manufacturers to actually cite those. And that won't be the case for all the claims because I understand you have to perform certain

studies with your own assay but in some situations might be applicable? Is that something that the FDA would consider?

DR. LIAS: Of course we would consider it and it really depends on the situation and the claims and the studies that you are referencing and how relevant they are. But for example for BNP or NT-pro-BNP there are many situations where some of the claims that manufacturers seek are supported by literature data and others are supported by clinical data on that assay. That has also in the past been true for troponin. So it really would depend on what the claim was and from our perspective it just has to be data that are relevant to the use and enough for the use. We are not seeking perfection. We're seeking good enough to outweigh -- benefits outweigh risk, that is really the decision point and or substantial equivalence if there is a precedent. But where you have no precedent and you are adding a claim you've got to just show us that the benefits outweigh the risk even if the benefits aren't 100% and if there are risks that is okay.

DR. KELM: What is not clear in that example

is whether or not the studies were done using your test or not. It may not matter but obviously if you -- if it then uses a different cut-off and it's not your device and that information would be you know put in a label in order to let people know how to interpret the result for that other claim if it is significantly different then how would you deal with that difference, the different assays being used, consideration or issue.

DR. GEE: So I have a question related to one of the topics that popped up this morning in terms of the rule out. Just curious would FDA be agreeable to a reanalysis of a pivotal data set that is used for the main aid in the diagnosis claim even if the primary end point wasn't to support a rule out claim. But would FDA be agreeable to a reanalysis of the data to see if a rule out claim would be supported?

DR. LIAS: So the general trial design wouldn't be different for rule out as long as the trial was conducted to get the early time points probably would be necessary. What's sometimes needed though is a different device design. So sometimes the cut-off

isn't appropriate for the values. However for troponin since the cut-off isn't usually built into the device it is usually in the label if one were to prespecify the analysis prior to doing it, you know we could certainly discuss how that might be able to be done.

MS. AJONGWEN: Based on all the information about expenses in doing the trial, if we have to take into account all the biases is FDA open to different methodology for example some sort of adaptive design methodology cross validation. What I'm saying is we keep on saying don't use the same data set that you set your cut point which make sense as a validation. But what about augmenting using some sort of adaptive approach to validate that cut point.

DR. LIAS: Sure we are familiar with proposals for cross validation or different types of splitting of sample sets. It is not ideal. But we have wonderful statisticians here who can have that discussion. So I think with any of these sort of proposals for using existing data or real-world evidence it is always a good idea to come and talk with us. Keep in mind we are open to this. We really want to be able to use

data if it exists. We don't want to make people go out and do a new study if they can use a study that is already there. I think the challenge people run into is often some of this data doesn't yet exist. And you know if we can either come up with ways to leverage data using valid statistical techniques that might be a good option. If we can figure out ways to design trials such that they can be used a lot of times and you get an optimal thing out of it, that might be also something good for the future. So I encourage all stakeholder PIs, manufacturers, physicians who are involved to come to us and talk with us about it. And don't assume that we wouldn't be open to it because what we want is what is good for patients and if some of these things can support uses that are good for patients we're very interested.

MS. AJONGWEN: It is nice to know because most people have been shying out of that because FDA come across like you cannot use the same data set that you used for your cut point. And I've already seen them think about doing some cross validation or mention that data to validate that cut point.

DR. LIAS: Well, you can't just -- we see people sometimes come in where they have looked at the data, set a cut point and then tried to use that data set exactly for --

MS. AJONGWEN: No, no.

DR. LIAS: You can't do that, that is what we are saying.

MS. AJONGWEN: No, that is not what I'm saying. I mean like using -- augmenting, adding on to that data to --

MS. LIAS: Right.

MS. AJONGWEN: -- increase your sample size.

MS. LIAS: The thing we are cautioning is not just diving in and using a bunch of things without pre-thinking about how you need to use it later.

MS. AJONGWEN: Yeah, yeah, I understand what you are saying. Okay. That is good to know. Thanks.

DR. LIAS: And I want to clarify there's very few things that FDA -- I mean there's a lot of statements today about what FDA requires here and what FDA requires there. There are very, very, very few things that we have as requirements for these. You

know we talk about what's good study design to support the intended use. There always may be alternate valid evidence to support a lot of these uses and sometimes those are practical to do and sometimes the trial we've been discussing an all comers trial is the most practical way to do it. But a lot of the things that you may think that we would or wouldn't accept may not be correct. So I would definitely come and talk to us and if you hear it from the horse's mouth that is one thing but if you hear it from other horses mouths you might not necessarily -- you might benefit from asking us again.

MS. ALVEY: Hi, Stacey Alvey. I think it is a great idea if we could potentially use existing data set or samples that exist because we feel that they meet our intended use population and serve our needs. Some of the feedback we've heard is that sample stability needs to be done with your investigational device. So if I'm interested in a sample set that has been in the freezer for two or three years but my investigational device is under development I don't have two or three years of sample stability with that

device. I'm never going to catch up to what is in the freezer. So we are interested to hear your thoughts on that.

DR. LIAS: So depending on the biomarker there's different approaches to doing so. In some cases you have biomarkers with a fair amount of data in general about its stability or instability. And if there is consistent data across methodologies or across commonly used methods or whatever to show that it is generally stable that type of information is helpful to understand. There are other biomarkers where there's almost no data on whether it is stable and there is some reason to believe it might not be stable because that type of molecule is often subject to degradation or things like that. And there are some cases where you know what the value ought to be from a different way of doing it and you can show that your device reasonably does that. And it is also never -- we don't usually get perfect data on some of these things. But we also -- so talk to us about what you do have and often we can figure out some way to figure out what's reasonable to show us stability. But sometimes we have

gotten things where we have 25-year-old specimens and they don't have any information at all to show us that that for example protein is stable for 25 years in the freezer.

MS. ALVEY: Can we use literature?

DR. LIAS: Yes, if it's literature available certainly. And a lot of times that is what is used and there are some analytes out there where stability is well established and this is not a question. There are other analytes where it is not established.

MS. ALVEY: How do you feel about troponin?

DR. LIAS: So troponin is an interesting one because the assays detect different epitopes and so what I would do is just think about what evidence is there in the literature. I think there is some evidence on troponin, there is some information about it. If for most assays it's got certain stability profiles, that's one thing. We do see some differences in sample stability based on handling across assays. So that is one thing you just have to think about a little bit about what your assay is and what it is similar to.

MS. ALVEY: Thank you.

DR. RICHARDS: You asked a question and just from industry I think some of the other issues that come up because trials are expensive and wanting to pursue additional claims is leaving MI out for a moment and thinking in terms of myocardial injury would FDA entertain the option of showing ROC curves instead of fixed cut-offs in labeling for other indications and that's again whether it is myth or truth, what's been out there is no, FDA wants a prespecified cut-off for any condition. But we do see other presentations of data that aren't necessarily based on a fixed cut-off. Can you comment on that?

DR. LIAS: I think we're happy to talk about different ways to present the data and why that might be helpful to labs because really at the end of the day we want information on the label that is helpful to laboratories for the way that they use these devices. And you've seen us put data at different troponin levels in labeling because in those cases it seemed okay. One of the things that in the past has been a challenge with the ROC curves is that entities may do one study and just present the ROC curve they got from

that study and there is no information to show that that is reproducible in other studies. So talking to us about one, how would a ROC curve be helpful on the label and two, how are your prespecifying some validation of that would be fine to discuss.

DR. NOWAK: You know I have a question. I guess it's for the FDA but it’s also for the clinicians. So in the trials that we've done if we are looking at a new troponin assay what generally happens is the treating physicians are blinded to that new troponin assay and what happens is the adjudicators either use a local hospital cut-off or they use a central lab cut-off but of a different troponin. And then what we do is we tally up the number of MIs that we've seen and then we try to fit in the sensitivity and specificity of the new troponin assay to what we've seen based on the adjudication process. And my question is is that the new troponin assay has a 99th percentile, it has a typical rise and fall in troponin one of which may be above the 99th percentile. So if you went back and actually took a new set of adjudicators and actually blacked out all the local hospital troponin Is and

looked at the number of MIs that were not diagnosed by the new troponin assay you would probably pick up new MIs that were smaller but not necessarily picked up by the older more contemporary assays. And if you did that what happens to labeling then. Because right now I mean I don't think the FDA looks at it but it is quite possible that there are people who have a small MI that is missed because they are using an older troponin assay and the new troponin assay is not given the same chance to make that diagnosis based on the third universal definition. So should people go back and actually look at trials and re-adjudicate them with the new assay and see how that re-adjudication process compares to the old adjudication process?

DR. LIAS: Are you talking about the trial that was used to look at performance of the new assay or a different --

DR. NOWAK: Well for example in --

DR. LIAS: -- new assay?

DR. NOWAK: -- TRAPID, we were involved in TRAPID. So that was a troponin I that the adjudicators used to make the gold standard and we looked at novel

ways to use troponin T for a one-hour rule in and early rule out and early rule in. We never went back and blinded a new set of adjudicators to the other troponin I assay but actually used the new assay which has a 99th percentile and looked at the number of people that would have been diagnosed as having an AMI but using the new novel one. It seems you should be able to do that.

DR. LIAS: So in a trial that is intended to evaluate the performance, clinical performance in this case of the investigational assay you wouldn't -- we wouldn't for our purposes of putting the label be able to assume that that was the right value because part of what we're looking at is the analytical performance characteristics and how that translates into clinical performance of the test. Now down the road as more of these new generation assays are on the market they will potentially be some of the ones used in the trial to compare to. This is an issue sort of that we talked about when we were talking about trials and what may happen in terms of the data. But it is a challenge not only for troponin but for other biomarkers where there

is some belief that the performance of a new biomarker is a little different than the performance of a comparator biomarker. So the purpose of adjudication in part is to say that in a scenario where you have additional clinical information beyond troponin you could adjudicate the whole clinical picture in order to help assess the performance of that new assay.

DR. NOWAK: And if you did re-adjudicate based on the new assay that you were testing you might get a lot more information on the analytics of it and actually look up the cases that were missed as small MIs. I just think it is an opportunity to see really what a rise and fall in a novel assay that actually has a 99th percentile determined what actually -- how that would perform and give you some more information on maybe whether it is actually better than the assay that you are testing it against.

DR. LIAS: In a different context that might be interesting. In the context of actually validating a new assay where analytical and clinical performance is unknown I don't think it would be helpful to us. However to a clinical community as time goes on to

understand the differences in troponin assays and how things evolve that would be good to know so that companies could make a choice about what comparators to use.

DR. SANDOVAL: I was just going to comment that just as your point that it can be done. It is actually what we did in our UTROPIA study with Fred Apple. So this was actually planned upfront design in which adjudicators adjudicated for both the contemporary and the high sensitivity assay and they were blinded to the other result. So we did it upfront for essentially both results. So we have -- this was a lot of work of course and that is the reason it doesn't happen all the time for most studies because it requires to go over each event for a whole set of different results but we have one whole set of adjudications using the contemporary assay, we have another whole different set using the high sensitivity assays. So it can be done but it was a planned upfront study design at least how we did it.

Well, we have a number of papers published on this but you know essentially I guess it depends on

what you are asking. We would look at this with the LoD. We've looked at this with the 99th percentile. We've looked at this with deltas. We've looked at this with ACG. So we have a number of different analysis and publications with this study.

DR. PILCHER: I'm going to interrupt you guys for a brief moment. You can maybe take this up later. We have one question that's been waiting a while and we are just about out of time. So if you can ask your question and --

MR. HUANG: This question is mainly for the FDA I guess. But this -- I realize it may fall under the come and talk to us and we can see and consider it but when it comes to using data from outside the U.S. I mean could you offer some general guidance. I realize the populations are more homogenous over there in Europe than in the U.S. but in terms of I've heard varying numbers from different sources. But could you offer some guidance on how much of the data, if at all, can be used from studies that are done with the same assay outside the U.S. as opposed to in the U.S.?

DR. KELM: Yep, so in many cases we accept

outside U.S. data and studies. We do often depending on the analyte or other issues ask some questions about anything from you know we do have some circumstances where we know analytes are sort of at levels internationally different here than other places and the practice of medicine may be different for -- those kinds of questions are things we always have. We do ask about demographics. There are ways that we could discuss trying to bridge that if that is a problem. But I think that we actually just have used outside U.S. data for troponin.

DR. LIAS: We actually have had some trouble using outside -- European data for troponin through it is not necessarily impossible. And depending on the claim for MI diagnosis. And because of the point that was brought up earlier about the way that they triage MIs and how the prevalence differs, the population differs a little bit in the centers over there. So that might not be true across Europe. So one of the main questions Kellie mentioned is do clinical practice -- are they similar and does that matter if they are not similar. So if you can talk to us about how where

you want to do the trial, how their practice is similar in terms of that and the demographics in the patient population --

MR. HUANG: Thank you very much.

DR. LIAS: -- and whether it matters. So I'll give an example on a different marker. We once got a vitamin D reference range study where they measured in Scotland and we asked them to do a little bit more in the U.S. because vitamin D reference ranges in Scotland are a little different than Arizona.

MR. HUANG: Thank you.

DR. SANDOVAL: Let me just comment to that about that use from outside the U.S. because the -- this was phrased briefly earlier but in the United States if you look at nationwide data there's a publication by Macom (ph) in JAMA I believe and this was also seen in our local data as well. So if you look at our study for example for UTROPIA it is I think 51% of the patients have chest pain. I think Korley from Hopkins had a similar like 56% of patients with chest pain. When you look at the vast amount of publications using high sensitivity troponin assays from Europe and

you look at the prevalence of chest pain in these populations, for example if you look at the meta-analysis from Chapman just recently published in JAMA for the non-U.S., non-North American sites the prevalence of chest pain is 89%. So there is quite significant differences. It doesn't impact at least to my interpretation it doesn't impact a lot the performance for ruling out for NPV and sensitivity but it does impact and influences the metrics for specificity and positive predictive value.

DR. PILCHER: I'd like to thank everybody on the panel and all the questions for the audience. We are out of time now and we'll move on to the next item on the agenda.

Thank you.

PUBLIC COMMENTS

DR. CAPOSINO: So we would like to open the floor up for public comments. I think we have three people who have registered to speak.

DR. SAENGER: Okay. Thank you everyone. So I'm Amy Saenger. I look a little different. I changed for the flight. But at the University of Minnesota. I

also am a member of the College of American Pathologists or CAP Chemistry Committee. And so I am here representing the CAP and giving some input and kind of our overview of what we see with troponin testing today in laboratories and kind of where we hope that we can help influence that.

So I'm excited to be able to talk about this because we meet three to four times a year, this committee and we talk a lot every time about troponin; like a couple of hours at least. Troponin and A1C we always talk forever about. And then some other stuff thrown in there.

So this committee what we do is we really look at ways we can incorporate improvements in proficiency testing based on new lab tests protocols. So laboratories are accredited by CAP as one of their proficiency testing providers and as of late 2014 in our cardiac marker proficiency testing survey we decided to introduce specimens which had very low concentrations because essentially we like to challenge laboratories and see how they perform throughout the range. And we didn't really have any specimens that

were kind of at or near the relevant 99th percentile knowing that we weren't going to get perfect values because of the lack of standardization or harmonization.

And so really I think most of you know the laboratories are required to report their proficiency testing results just as they would patient results. And so they have absolute, less than, and also greater than values can be reported in these schemes.

And so one of the things that we noted looking at the results among peer groups within these low concentration samples what is shown is the number and the percent in parenthesis of laboratories who are actually reporting values that were below the limit of detection. And these are all with contemporary assays. And so what we found is there was really a wide array of values that laboratories were reporting, a lot of them were reporting down to zero which isn't a real number. Some of them had a range of values, the lowest reportable is .2 up to -- there isn't a lot of consistency today.

And so I think that is something that we hope

to improve with high sensitivity assays or at least give some guidance on in addition in conjunction with the AACC Academy and IFCC Task Force.

And one of the things that we know that has been coming is that they recommend reporting these values, concentrations in whole numbers for high sensitivity assays in nanograms per liter.

So we kind of thought we have mostly U.S. laboratories but we also have a fair number of international laboratories that participate in our proficiency testing schemes. And so we kind of thought as laboratories are introducing high sensitivity assays we'll kind of see a shift in how laboratories are reporting. And it would be "obvious" that laboratories are making the shift and we can kind of see and gauge where we're at.

When we looked at -- this is data from last spring and not surprisingly most of the U.S. participants report some form of troponin I in nanograms per ml. And there were a fair number of U.S. participants that reported both INT in nanograms per liter, so I am not really sure, this was pre-

introduction of any high sensitivity assay. So we were kind of surprised to see that.

But when we looked at just the international participants for high sensitivity cardiac troponin T amongst the different platforms surprisingly what we saw, we thought everybody would be reporting in nanograms per liter because the recommendations have been out there for quite some time to report in whole numbers and nanograms per liter, what we found was really that there was a big split in how laboratories were reporting internationally and it was split amongst a whole host of countries. So we couldn't just say oh, it was the Canadians who were doing the wrong thing.

And so I think and we found actually the same thing for I there is a lot less participants because there is a lot less assays that are CE marked. But we found a similar thing particularly for the Abbott high sensitivity troponin I, a large number of participants were self-reporting using that designated high sensitivity assay in nanograms per ml. And so we really feel that there is an opportunity with the introduction of high sensitivity assays here to promote

the use of nanograms per liter to report in whole numbers. I think now as CE in the U.S. there's options available for labs to choose from as to how they are going to report even in nanograms per ml or nanograms per liter. So I don't know if there is a way to help standardize that on the labeling. But we kind of see that if given the chance people will kind of just 50/50 pick one way to report or not.

We also did a survey with our last proficiency testing samples and we actually had a good number of U.S. laboratories who responded over 2500 labs and we just basically asked them kind of what you reported as your abnormal cut-off and gave them a whole host of choices to choose from. So most of them or 38% I guess said that they used the 99th percentile. Some didn't know, some used other which I'm sure what that might mean. Some used a literature based cut-off. Derived their own. So there is a lot of differences in how we're reporting troponin today. We'd like to kind of standardize so that everyone is at least ideally reporting or using the 99th percentile to flag abnormal results.

And actually when we looked at the individuals who checked the box that said they were using the 99th percentile and then we asked them to actually give us a number the range of kind of what they said their 99th percentile was was quite different. So for some assays the Roche contemporary troponin T some said they used the 99th percentile and it was less than .01. Some said they used the 99th percentile and said it was .1. So I don't know the labeling is a little bit sometimes difficult to understand with the current assays as to what cut-off laboratories should be reporting. So our thought is hopefully with the new assays coming onboard that we can clearly have consistent terminology and verbiage and guidance to labs as to what to use.

And we also asked if they used intermediate or grey zones, this is another thing that is kind of inherent probably from the last -- for at least the last decade and for those that responded that they did use intermediate or grey zones which means maybe somewhere between the 99th percentile and maybe the WHO cut-off and most of the time they weren't flagging until the WHO cut-off. Most of the laboratories that

responded yes that they do use a grey zone were in the U.S. So we still feel like there is a lot of education that we need to do in this area.

And then finally I'll just say we asked a question about turnaround time, goals and metrics. And the previous recommendations from the NACB which is now the AACC Academy was that the turnaround time should be less than 60 minutes from the time of blood collection to reporting of results. And we asked laboratories actually how they collected their turnaround time data and how they defined their metrics. So a majority of laboratories actually collect turnaround time data from the time that they receive the specimen in the laboratory to the time they result it in the op instrument. And then specimen collect report there were less users, most of those, about 35%, were point of care users. And so I think a majority of labs are able to control the tracking specimen turnaround time from the time it hits the door out. But when you look at a lot of the rule out algorithms and from the clinical perspective how they are actually defining a turnaround time of a baseline zero or one-hour, two-

hour strategy is from the time the patient hits the door of course because that is what they are seeing. But what we are seeing on our end is something that is a little bit once it hits our laboratory.

So you know the discussion could be some of the turnaround time requirements definitions, they do differ between specialties and in the laboratory. I know Dr. Apple will probably discuss this but the new recommendation will be less than 60 minutes from the time a specimen receipt to reporting results.

So these are some of the quality metrics that we hope to also encourage laboratories to report and share consistently with their emergency departments.

And to conclude I just wanted to reiterate essentially that currently we don’t have an acceptable way to routinely evaluate the performance of current assays. With high sensitivity assays what we plan to do is implement a plasma serum pool into our proficiency testing scheme, probably a male and a female kind of normal pool which we'll send out with all of our proficiency testing samples. And that will give us a real sense of how all the different platforms

are performing across the U.S.

In terms of reference intervals reporting unit turnaround times we feel like there is a huge opportunity to standardize and educate and we've also talked about having separate checklist questions available for when laboratories have their I guess biannual CAP laboratory inspections. They'll be kind of required to show validation data, verification data as to where they get their cut-offs. And also a reinforcement of the acceptable reporting limits. So if it is a checklist question essentially the labs kind of make the changes happen. If it is not a checklist question then people say oh, that's nice, it is in the guidelines but I don't really have to do it. So we feel like this is a way that we can at least help give specific guidance to a large number of laboratories.

And finally I'll just say I personally think of the joint relationship between the emergency department, cardiology and laboratory medicine as really this three-legged stool and so we have a huge opportunity to as we move forward to work together. And if one's leg breaks it all kind of falls down. So

that is all I have and wanted to comment on.

DR. CAPOSINO: Thank you.

[APPLAUSE.]

DR. CHRISTENSON: Amy. Amy just a quick question. Has CAP ever thought about a recommendation of running a control near the 99th percentile, I mean near kind of where the money is rather than so many --

DR. SAENGER: Yeah.

DR. CHRISTENSON: -- assays, run them way up here.

DR. SAENGER: So we have that low sample which is supposed to be absent of troponin. The problem is is that more with troponin I is that you can't get a sample that is like low across all the platforms and assays. So that is why at least some will just give you undetectable or less than a value which won't be useful to that peer group. So we feel like once there is more high sensitivity assays onboard we'll be able to have the kind of normal donor serum or plasma pool, send them out and see how the actual instruments are performing in real time across different laboratories. But right now we are kind of stuck.

DR. CAPOSINO: Okay. There's a ladybug up here and it made it up to the microphone.

Our second speaker. All right. We'll take you.

DR. APPLE: Thanks for the opportunity. Fred Apple. So I'm going to wear my hat now as the chair of the IFCC Task Force for Clinical Applications of Cardiac Biomarkers. At a request of consideration for the FDA that we published a couple of years ago that the definition of a high sensitivity assay. So my request from my task force if the FDA consider if they are going to consider designating assays as high sensitivity they use that terminology that our international expert opinion task force has recommended and use the criteria that we have proposed in publication and peer reviewed literature. It is endorsed by the Journal of Clinical Chemistry, it is endorsed by the Global Task Force for the universal definition both by the third and soon to be fourth universal definition, and it is endorsed by an expert opinion group from the AACC Academy.

So we just put that forth to you to consider

that if you designate because it's confusing in the literature if and I pick on manufacturers that they publish names of their assays ultra this or super sensitive and we feel it is important to have uniformity globally. If you are going to be part of the global world and designate assays maybe you won't but if you do high sensitivity is the expert opinion terminology we have endorsed.

Thank you.

DR. LIAS: So unless there are anyone else who wants to make a public comment I'm going to give a brief summary and hopefully talk a little bit about what I hope to be next steps.

Before I do that Fred actually we have no objection to people using the term high sensitivity. All that we've asked of manufacturers is that if they do that they define what they mean by it. So for example if they meet the definition that IFCC had designated that they can say that they are high sensitivity per that definition. So we don't have any problem with them doing that.

CLOSING REMARKS

DR. LIAS: So I want to personally thank everybody for their participation in this meeting today. You know we had some pretty I think ambitious goals for today in terms of covering a lot of topics. But we really wanted to do and I hope we at least partially accomplished was start to open up some lines of communication.

I think that there are a lot of people working in the space on all sides. There are manufacturers who make troponin assays. There are laboratorians who run these assays. And there are doctors of various specialties who use these assays. And everyone has different perspectives. And for us all to try to understand the different perspectives and try to come to some understanding of the different ways that troponin might be used, might be studied, and might be made available to meet all of the needs for people and I think that is the ultimate goal here.

And to do that we all need to talk together and we all need to openly discuss the challenges that we have and how we might solve them and what we wanted today was to really start that process off.

There are several things that we learned today. This morning we talked a lot about reference range studies. And I think what was clear to me is that there is a lot of agreement but there is still some disagreement about how to do some of these studies in the right way. And I think one of the parts that maybe isn't well known is exactly what population is the population that should be used to define a reference population since the reference population for troponin is being used to determine clinical cut-offs whether you like that or not. But that is necessarily being used right now.

And so perhaps some work remains for us to continue to talk more as a clinical community to determine what are some recommendations for maybe getting some more harmony or standardization among how these cut-offs are developed so that the clinical community can really understand across assays what they might see.

We talked a lot about analytical issues and pre-analytical issues and how consideration of those issues prior to doing clinical studies can really

prevent some of the challenges that manufacturers have seen in clinical studies and give laboratories assays that can be labeled such that they know how to use them.

We talked a lot about what types of claims might be useful to clinicians to use for MI, that maybe differentiation between type-1 and type-2 MI isn't so great. But perhaps the addition of devices designed to rule out MI might be beneficial to everyone.

And it seems like one of our next steps might be to all work together to figure out how do we get to that stage. How do we encourage the development of devices that can do rule out? How can we come to agreement on what might be acceptable performance for a rule out type device?

We talked a little bit about point of care devices. And heard some perspectives that perhaps for these types of high impact claims point of care devices need to work well as well as they need to work which in some cases might be just as well as a laboratory device. In other cases if they are doing a different purpose maybe there could be some tradeoffs. But that

that needs to be thought through and they need to be robust to the environment that they are used in.

And finally we talked about ways we might be able to make trials more efficient, to leverage data that's existing. And perhaps to bolster our infrastructure so that in the future we might be able to develop this data more quickly.

So for me my personal take home messages were many. One, I got to connect with a lot of people today that I hadn't been able to talk to in a while. And I hope to continue those conversations so that we may move some of these discussions forward.

I think as a next step we as a community need to decide where more clarity is needed. And I think what is clear to me is that we at FDA have not done a good enough job of giving our perspective on things such that we might dispel rumors of what would or wouldn't be acceptable. We need to make sure people understand that we want what is good for patients and what's good for doctors and we need to be able to give doctors the tools that they need to do the work that they have. How can we provide an environment that is

conducive to having people feel comfortable about talking to us about how to get that done?

And so hopefully we've made those strides today.

And I want to thank you all for being present in this conversation.

We're certainly going to take this information back within FDA and talk about whether additional conversations moving forward on some more focused topics might help move this forward. And we are always open to suggestions on where you all think the most productive conversations might happen.

So once again I'd like to really thank everyone for coming. And I hope you have safe travels.

Talk again soon.

Thank you.

[APPLAUSE.]

CERTIFICATE OF NOTARY PUBLIC

I, NATALIA THOMAS, the officer before whom the foregoing proceeding was taken, do hereby certify that the proceedings were recorded by me and thereafter reduced to typewriting under my direction; that said proceedings are a true and accurate record to the best of my knowledge, skills, and ability; that I am neither counsel for, related to, nor employed by any of the parties to the action in which this 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.

NATALIA THOMAS

Notary Public in and for the

State of Maryland

CERTIFICATE OF TRANSCRIBER

I, CHERYL LaSELLE, do hereby certify that this transcript was prepared from audio to the best of my ability.

I am neither counsel for, related to, nor employed by any of the parties to this action, nor financially or otherwise interested in the outcome of this action.

12/11/2017 _____

DATE CHERYL LaSELLE

1

2

3

4

5

6