(Intro- Greg Whitney)
:15 Okay, hello, everyone, and welcome to today's iteration of our CIL (Cambridge Isotope Laboratories) hosted Webinar Series.
Delighted to have everyone with us today, as always. Certainly, in these sessions, we endeavor to dig into some of the exciting and relevant content from the wide world of stable Isotope applications.

:33 And delighted to have Huafen Liu with us today. She is the founder and president of Calibra Diagnostics. I will give her a proper introduction in just a moment.

:45 But she is going to take us through today the Opportunities and Challenges of Metabolomics Biomarker Discovery and Development in Clinical Diagnostics.

:53 Just by way of brief introduction, my name is Greg Whitney. I am our Global Sales Director here at Cambridge Isotope Labs, and once again, delighted to have all our participants with us today.

1:07 We recognize everyone has busy schedules, of course, and delighted you could spend few minutes today with us. So, in brief, just a few goals for today. Obviously, we want to help everyone explore some cutting-edge applications of stable isotopes. We hope you learn something new and something practical from today's expert speaker that you can apply in your own laboratory.

1:29 We would very much like to strengthen your understanding of the role that CIL and our products play in these 'Omics research areas. And certainly, we want you to stay engaged to have some fun. We recognize there are lots of commitments on everyone's calendars and schedules. I'm so happy to have you with us today.

1:50 From a format standpoint I have just a few quick comments and a few questions, just to help us get a better sense of our audience today.

2:00 I will then hand it over to Huafen, our keynote speaker for the day, and then Andrew Percy, who is our senior applications chemist and our MS (Mass Spectrometry) 'Omics product manager here at CIL.
He is going to help us moderate our Q&A (Question & Answer) session at the conclusion of today's session.

2:18 Just a few administrative reminders. All our participants are going to be muted for today's presentation. We will be recording today's event and distributing it upon conclusion. We would very much encourage you to utilize the Q&A functionality in the Zoom platform to submit your observations or questions for Huafen during the presentation we will address all of those in moderated fashion at the end. And I would also remind folks that closed captioning has been enabled for today's talk as well.

2:55 Okay, so just a couple of comments from my side for those of you that may be new to CIL. We hope, certainly, that we can be, connecting with and reaching net new contacts through these types of events.
3:09 But at our core CIL is the world's leading stable isotope company, and we endeavor every day to demonstrate our expertise in stable isotopes and organic chemistry. And some of the hallmarks of that expertise includes full scale production capacity. So, we can work with very small, precise quantities of material all the way up to multi-ton production capabilities.

3:35 Of course, these products are used in a multitude of applications. Several key areas I have listed here, but our product portfolio is comprised of some 15,000 different products, that aim again to help you in a known wide variety of research areas and production.

3:57 We are very proud to be the preferred supplier to analytical laboratories worldwide and while we have extensive direct sales presence domestically and internationally, we also have a very expansive international distribution network to put our products where our customers need them when you need them.

4:16 These sites comprise the CIL group. Our corporate headquarters is here in Tewksbury, Massachusetts, which is where I am based, here today.

4:28 We also have a production facility just a couple of miles down the road, in Andover Massachusetts, where we do a lot of our wet chemistry, again, small scale to large scale.
We also have GMP production capabilities in both Tewksbury and Andover. Our site is in Xenia, Ohio which is where we do a lot of our large-scale isotope separation of both 13C and 18O from natural abundance.

4:55 We also operate one of the world's largest D_2O re-enrichment facilities which helps support our self-sufficiency in a very substantial way.

5:06 Our site in Canada, in Montreal is where we do a lot of our biochemistry, primarily production for carbohydrates and amino acids.

5:15 Our 2 subsidiaries in Europe: Eurisotop is just outside of Paris, France, as in many ways, some similar capabilities to the work that is done in Andover, and our facility ABX in Dresden, is the only site within the CIL network that works with radio isotopes primarily for products related to nuclear medicine.

5:37 We also have sales and marketing presence in Shanghai China as well. And without going through every specific element of the organization for the products and the applications that we will be discussing today, our site in Andover is probably the primary site where a lot of this work in production would be done.

6:01 And again, just for background's sake. This is where we would do a lot of our synthesis, purification, formulation. Again, all done on site here at CIL, in Massachusetts, and we also offer

a wide variety of custom synthesis and custom formulation, which can help you with any needs for unique standards or multi-component mixes that might support specific elements of your research. And with our obviously robust quality control in process QC (Quality Control) testing, you know the this is really a hallmark that we would obviously you know, lean on day in and day up for our customers.

6:42 We have a wide variety of analytical techniques that we utilize to make sure that products that we provide to you are reliable, that meet specifications and are of the highest possible quality.

6:55 Okay, so with that we do like to take this opportunity before we dive into our content to try to understand a little bit more about our participants today. So, we have just a few questions that my colleagues will help me bring up and will post them to our group. And hopefully, as you guys are sort of multi-tasking and getting through today's content if you can help us just for a moment with these questions. It really gives us a tremendous understanding of the work that you are doing as chemists and participants today, and it is very helpful to us.

7:39 So question number one is: how many of our participants today are currently performing quantitative or qualitative mass spectrometry? And we would encourage you to check all that apply, either one, both, or neither as the case may be.

7:57 The second question helps us understand the relative frequency with which you are utilizing stable isotopes in your research and in simplest terms, if you are not using stable isotopes at all. Perhaps you use them sometimes, or infrequently or in some cases, you are applying them in daily use within the context of your research. And then last, but not excuse me.

8:27 The third question again, just from supplier standpoint, it is helpful for us to know that when you have needs for stable isotopes it is helpful to us to understand is CIL your preferred supplier at this point, and hopefully, events like today, and you know additional content that that we put forth hopefully that helps support
your perspectives of CIL as a source, moving forward.

9:00 And question number 4, which I believe is our last one. It would be helpful for us to understand. if you have used any of the following CIL products. Our Credential E coli Cell mixes, our Metabolomics Yeast Extracts, our Metabolic QC (Quality Control) Mixes or kits, our TCA (tricarboxylic acid) cycle mixes, fatty acids or lipid standards, nucleic acid standards or quantitative metabolomics mixes and kits, which we offer a variety of for amino acids and organic acids.

9:36 So, I am very grateful for those that have taken a moment to indicate their responses. We will give everyone just a moment to fill in their final answers, and then Crissy will conclude the polling for this morning, and then I will look forward to handing it over to Huafen for our main content, for today.

10:18 Okay, thank you too all. Crissy, if you would not mind publishing or posting the responses that would be excellent. Okay. So just by way of a very brief rundown, it does look like we have several participants doing both qualitative and quantitative analysis, which is good. As we would expect most of our participants today, utilizing stable isotopes either always or sometimes, which is again to some extent to be expected for today's content.
but helpful to know. Also, interesting, and helpful and supportive to know that CIL is a preferred supplier for many of our participants today, and certainly for those where we are not currently a preferred supplier, we are hopeful that content like today's can help us strengthen that position in your in your laboratory.

11:17 And also interestingly and very positively great to see kind of a smattering widespread usage of a variety of different products, most notably some of the quantitative mixes and kits. That is great news, and we are excited that our participants are putting those products to good use in their laboratories.

11:42 So thank you very, very much for that. And without further ado, I am delighted to hand over the presentation to our keynote speaker today, Huafen Liu, who is the founder and President of Calibra Diagnostics which is a leading 'omics and clinical diagnostic total solution provider in China. Calibra does focus on the application reagent and product development of a variety of mass spectrometry applications in clinical diagnostics and the life sciences.

12:20 Huafen previously worked at Pfizer, Roche, Sciex in the United States for more than 20 years. She was the Global Director of Sciex before founding Calibra in 2017. She has published more than 100 different publications and received the AOAC International Award for technical and scientific excellence in the 2015-2016, period. And today she is going to take us through the opportunities and challenges of Metabolomics Biomarkers Discovery and Development in Clinical Diagnostics.

12:54 So with that Huafen, thank you very much for joining us, and I will happily hand the presentation over to you to take it away.

Huafen: Thank you, Greg. Let me share the presentation. Can you see my slide?

Greg: Yes, we can now, You're in presentation mode.
That is perfect. Thank you, Huafen.

Presentation (Huafen Liu)
13:28 Okay, great thank you very much Greg for your introduction. Hello, everyone! My name is Huafen Liu. Thanks Cambridge Isotope for inviting me to give this presentation.

13:40 So the topic, as Gregory introduced is the Opportunity and the Challenge of a Metabolomics Biomarker Discovery and Development in clinical diagnostic settings. In today's presentation I am going to present the following content, including a brief introduction: The key challenges and the opportunities and the importance of reliable data and this is one example of a Biomarker discovery and product development.

14:13 And finally I am going to share some regulation requirements for the *in vitro* diagnostic multivariate assays.

14:27 So 'omics, including genetics, proteomics, and metabolomics represent different biologic functions as the audience knows. DNA is the genetics information carrier which should tell us the potential risk of the diseases from transcription the protein is synthesized.

14:53 Protein is the functional executer in our body.

14:57 When the protein is functional and generates correlated endogenous, biochemical, small molecules, these are metabolized.

15:08 Metabolomics actually, it is the phenotyping and the closing to the disease.
Like the clinical setting which could be impacted by factors, such as lifestyle, aging and the environmental changes and the other factors.

15:25 That's part reason metabolomics there's multiple dimension factors. It is challenged in terms of study and research.

15:40 Dr. Leroy Hood, who is the founder of the Institute of a System Biology, and in year 2017 he contacts a study, a wellness study for 108 individuals using a personal dense dynamic data cloud, as shown that here it is including gm protein and metabolites and the clinical laboratory settings, and microbiome.

16:11 In this study it indicated that the metabolome provides breach data reporting on genetic risk the microbiome, and the clinical laboratory measurement, and the numerous metabolite correlation with the clinical labs suggests that metabolomic information can come complement and enhance the clinical practice and the understanding of individual health and disease.

16:45 Mostly importantly, it is also indicated that the genetic trees determine the risk that may or may not actually materialize, even materialized risk may be able to be managed by using personal health data cloud which is impacted by other factors in our lifestyle.

17:11 Life, it is a complex biochemistry process of which is the most phenotyping is metabolomics.

17:23 Current accurate mass system technology enables us to full scan, broad mass range to identify and quantify biological chemicals in the samples.

17:36 We can discover the biomarkers by comparing the normal population and disease population to guide the insight of the biological process and understand and search out the biomarkers. Even if all the values and the opportunities that metabolomics could provide to us, however, we are still facing many challenges.

18:08 Biomarker discovery is a systematic task. It is needed to join the effort and the integration of all kinds of different resources.

18:20 First, we need the physician or medical specialist to understand the disease mechanism, and we also need to find the right targeted population.

18:34 And also, we need to think about if there is any alternative method is using for diagnostic.

18:44 And for biomarker discovery it is deep research. First is most important thing being the cohort's selection and a sample collection, because without the fundamental, the good sample cohort we could now get the valuable information.

19:02 The second thing is the project design, there is a good statistical power.

19:08 Otherwise the variance finds out, or the biomarkers, we find out it could be misleading. And high quality is the cornerstone for the biomarker discovery. And, with such a large amount of data we need half power tool for the data mining and statistical analysis.

19:39 And also, we need correlate those findings associated with biological pathway to understand if those findings are truly biomarkers.

19:47 Finally is the productization. During the production process it is quite a challenge, and a long process. Sometime even, we find a good biomarkers are not necessary.

20:06 It can be turned into a product because we need to understand if a pure reference standard is available, and the compound is whether it is stable in the specimen, and whether it is applicable and repeatable with a different methodology and the QC (Quality Control) and linearity range for quantitation.

20:30 Most importantly, if it is a new marker, we need to establish the clinical reference range and the algorithm and need to relate to a clear clinical indication.

20:46 And also eventually from the commercial aspect, we need to consider which instrument we should choose, and the throughput, whether it is enough to support the need, and whether it is cost efficient enough for the patient to pay.

21:05 And then for the IVD we need a thorough validation and need to go through the registration requirements.

21:14 A meaningful cohort design is the cornerstone for the biomarker discovery.
First, we need to understand if currently we have any other golden standard available to compare as a reference point. And the control group of selection is critical. We need to address, use the group control group to address whether it is unique for this disease. And we need to understand the progress of disease to identify different types of disease.

21:57 And we also need to have enough samples for statistic power.

22:04 At the same time, a lot of people may miss the sample matrix choice, and the storage condition.

22:20 Because if during the shipment or any situation, if the sample is compromised, the data is not reliable, and the most critical thing, it is the sample collection needs to be from multiple different centers to minimize the bias, and this is systematic errors.

22:42 Here is just a typical discovery and development metabolomic biomarkers workflow. Typically, the disease population must match the age, gender, and other clinical chemistry of similar control group. And the first step for a totally new discovery we need first untargeted.

23:12 Typically for the samples we choose more than 1,000 from different multiple centers.

23:18 This is as a biomarker discovery 'omics study typically, we need to have more samples for discovery stage and from multiple centers to ensure that, the findings are true and repeatable.

23:38 And then, second, once we get these untargeted data, we need to use statistics and bioinformatics to choose most relevant the top 10 to 20 biomarkers and then develop the targeted assay.

23:58 In the targeted assay it needs to be quantitative. In this situation, I need very thorough message validation, good reference standard, and the internal standard.

24:17 And eventually the final target biomarkers better choice is less than 5 of them, because when you have multiple variants, it is really challenging to conduct the IVD registration for clinical trial and, we need to conduct multiple centers, and to ensure that in the confirmation stage, those markers are repeatable.

24:49 And eventually the last step is the first we developed these Assay in LDT in the central lab settings, and as a test-out, and then polish the message eventually to do the IVD registration. For a screening method in China, we need all more than 50,000 samples for our quantitative method we need more than 1000 samples from 2 to 3 different sites.

25:26 Reliable data is the cornerstone for the biomarker discovery. Here I am going to share some examples of a biomarker discovery, and the product developmental process.

25:41 First, the for the untargeted study. Good QC (Quality Control) is the foundation to generate reliable data, as we all we made familiar that for untargeted screening with the accurate mass system, we can get more than 10 thousand ions in one scan and however, how are we going to identify these are true biomarkers related to the disease.

26:14 We need to have a good QC (Quality Control) system with randomization in the sample process and sample run.

26:24 This is well to ensure to QC the data quality during the sample collection, and the data collection to minimize the bear variabilities during this process.

26:40 Typically, we gave more than 30 internal standards for each untargeted run and the current we were able to keep the instrument variability around 5%. At the same time in order to make sure the biological, the QC System we pooled the samples together, and use those as endogenous biochemicals QC.

27:12 These variables keeps within 10%. During the sample run we randomize QC and different samples.

27:21 In that way we can minimize the system variability.

27:29 Here is a good example why standard, and the internal standard is important to identify as unknown.

27:39 This is a just example of the database in our system, and the sample collection, as we see at the bottom.

27:50 As you can see over here, we compare the use accurate mass system. This mass accuracy less 3 ppm and this matching of the MS2 data and using the sample match the library and, we have the third level confirmation is from the retention time.

28:11 So this way we confirm the compounds are what we identified.

28:20 The other thing is that isotopic internal standard.

28:26 It is really critical for the untargeted analysis. As you can see over here, in our untargeted, we have more than 30 stable isotopic labels. The majority was purchased from Cambridge Isotope.

28:41 And this is a positive ionization method. We have the RSV for internal standard is less than 10%. And for the negative mode, it is less than 5% most of the time.

28:57. The other thing is, we need to keep in mind the repeatability of the biomarker discovery.

It is critical. It will ensure us what we find is true.

29:12 And typically we have different test batches of sample analysis. This is just an example of what is to be done about cardiovascular disease. As you can see here for the first sample set, We have a total of 60 samples. We find we identify more than 827 metabolites, and then the second batch of the sample we have 886 metabolites identified.

29:48 And among all of these, the top 10 most important candidates potentially could be developed as a biomarker. Several of them were repeatable which, indicating these 7 are reliable biomarkers that we can carry through the development.

30:09 Here I am going to share some proteomics and metabolomics study we have done during Covid-19. This study was carried out by Calibra Diagnostic, Taizhou Hospital in China, and West Lake University.

30:33 This sample set was collected at the beginning of COVID-19 back in February 2020. As you can see over, here this was the first publication for proteomics and the metabolomics characterization of the Covid-19 patient serum sample.

31:01 And actually, during this study we screened a series of potential biomarkers in the serum, and then for the Covid-19 patient and they build a machine learning model to predict the risk of a progression, from non-severe to severe Covid-19. And that this was this paper was also commented by Dr. Francis Collins, as the first molecule level profile for Covid-19 patient.

31:38 Following that, in the year 2021 we publish another paper from this same sample-set comparing the urine biomarkers and the sera biomarkers.

31:58 The study was designed, initially, this total population 118 samples.

32:07 This includes 28 healthy subjects and most importantly, we design a control group. It is a non-Covid-19 with a similar fever and the lung infection population but Covid PCR (Polymerase Chain Reaction) testing is negative.

32:24 This is critical to identify whether the Covid-19 patient symptom, or the biomarkers is specific, for Covid-19 and then non-severe 37 people and then 28 severe sample.
Matching age, BMI, medications chronicle disease, smoking, alcohol, and other lifestyles.

33:05 And for all these 118 samples we obtained 1,494 proteins from the serum, and 904 metabolites.

33:15 And in the urine sample, we obtained 3,854 proteins and over 1000 metabolites.

33:26 And using the statistic, including random forests we can classify and predict severe versus non-severe.

33:38 This classification accuracy over 85%. As you can see on the right-hand side the bottom and using the in the metabolite, simply, we can see the severe and non-severe
Covid 19 patient is very much different from the non -Covid-19 with fever and the lung infected patient.

34:05 And as you can see, those markers it is unique for Covid 19 infection.

34:14 In this study we used a machine learning-based classifier to identify the severe versus non-severe.

34:25 In the final model we chose, it's including 7 metabolites and 22 proteins. As you can see, there is a training set is 18 non-severe patients, and 13 severe patients.

34:43 And combine the proteomics and the metabolomics biomarkers using machine learning, and evaluate.

34:52 The first test cohort was 10 population and the accuracy, and the AUC was 70%, and then the second cohort was 19 patients, and the AUC was 84.2% as you can see here.

35:15 Covid-19 changed everybody's life, including mine.

35:22 And we also look at other applications in these special conditions like the gestational diabetes mellitus (GDM) study.

35:41, it is another good example. The GDM is a type of diabetes that you can develop during the pregnancy for normal women when they were not pregnant, and in the US, it is around 2 to 10% of the pregnant women has GDM.

36:06 And in China the percentage is even higher, it is up to 20%.

36:13 The GDM can impact on the infant like eventually the baby could get obesity, or other diseases.

36:26 And for the mom, there are a lot of risks as well. It could be spontaneous abortion, and/or other diseases.

36:40 The GDM study current is a Gold Standard. There are 2 different strategies.

36:47 The first one it is one step the OGTT tolerance it is going to take a 3 time point: the fasting, and then 1 hour after drink the sugar, and then 2 hours after drinking the sugar.

37:06 Another type for confirmation is the 2-step strategy. First it is performed for 50 grams of non-fasting at the glucose tolerance, and then the second step, using the 100 gram OGTT glucose tolerance.

37:26 So in the confirmation at its total 4 time points for sample collection, and the women need to wait in the hospital for at least a half day. During the Covid-19 pandemic stage, different countries have different approaches from Australia and Japan and the England Society of Diabetes and the pregnancy they have updated new guidance for the screening and diagnostic in GDM. And, during the conduct, the pandemic stage.
And the other authority suggesting, instead of waiting in the hospital, because it is going to get the risk for the pregnant women to be infected, so in that situation using the HbA1c or other factors to be an indicator for GDM.

38:37 However, in February 2021 in Paris 7 University with its hospital has published a paper indicating that HbA1c replaces all glucose tolerance, and it is not reliable, and this is a low sensitivity.

39:06 We think metabolomics and the metabolite pathway could answer this question.

39:15, This work has been conducted for more than 2 years. First initially we develop a biomarker. Our goal is to develop a biomarker to replace OGTT for diagnosing GDM.

39:32 And instead of drinking the sugar and we are thinking about just one-fasted sampling, and this is non-targeted metabolomics to identify the upstream and the downstream of the enzyme, and this see multiple pathway changes and to diagnose GDM.

37:58 The study was designed as a following for the targeted part. For the untargeted part I want to repeat it over here, since I have a pretty good description up front for the regular untargeted flow.

40:13 So we identify more than 15 biomarkers, and then we collect another cohort for the targeted approach. In this target validation stage, we collected 499 samples in total, including the GDM positive 97, and GDM negative 402.

40:42 And this GDM confirmation was using the OGTT tolerance experiment as a Gold Standard, and then we collect the age, and the pregnancy, BMI, and the clinical data, including the OGTT FPG and the one-hour glucose and 2-hour, and the HbA1c.

41:12 And for the non-targeted metabolomics was performed early on and this is here, we are using those 15 analytes as a target.

41:25 And we combine these and oh, it is a typo over here.

41:36 It is 19 different metabolites were detected, and 9 of them were a combined with a clinical indicator, were used to build the machine learning model. This includes the BMI and also the C Peptide.

41:55 Here is the statistical result. This is machine learning, and this total 13 variables eventually for the random forest of classification, and other statistic to eventually we used 13 variables to build that this model. And the EUC is point .86.

42:23 The specificity is 62%. The sensitivity is 94%.

42:30 The false negative rate is 6%, and the false positive rate is 38%. As you may see over here maybe the specificity, or like the false positive rate, is not optimal.

42:48 However, this is a risk assessment of 2.

42:54 We intend to set the benchmark this low, false negative rate because for the pregnant women for just one blotter tool, if they still have a high risk, and then they can go to next step for the confirmation OGTT Test.

43:23 So this way it will minimize the risk for false positive.

43:29 And the regression model result as the indicated over here, using the 0-hour fasting glucose level, we can, and the metabolite and we can predict 1 hour and 2-hour OGTT result.

43:54 This correlation for 1 hour is .96 and that this for 2-hour the correlation is .93.

44:07 And for this project it is ongoing progress, and that we are going to narrow down to a smaller set of analytes for clinical trial for the regulatory requirement IVD multiple variance assay.

44:33 In both US and China, FDA (Food and Drug Administration) has a certain regulation, for *in vitro* diagnostic Multivariate Assay, which is also called as IVDMIA.

44:49 For FDA (Food and Drug Administration) It has clear definition that a device that the combined values with multiple variables using an interpretation function to yield a single patient-specific result.

45:09 This is called IVDMIA. So, in China it has similar indication and for China for the screening assay is required more than 50K samples for a quantitation diagnostic and need at least more than 1,000 samples for the clinical trial for a Class 2 or Class 3 registration.

45:39 During the IVD registration FDA (Food and Drug Administration) also consider whether for this is the specific reagent, or the assay. Compared to other assay whether it is has its own special value.

46:03 Typically it is going to be asked is there any effective treatment for the positive result. Whether this diagnosis is valuable for medical decision.

46:14 The other thing is false-positive, and the false-negative consequence, just like the OGTT example I shared previously.

46:25 For a screening assay typically, we intend to keep false positive rate high, with a very low false negative to make sure that the high-risk population is identified.

46:44 And also, we need to minimize what is the minimal acceptable method requirement.

46:52 The other thing is that needs to be considered. Is there any other similar product on the market?

46:57 What is the new product value bring to this diagnostic?

47:04 Say, for example, the OGTT Diagnostic is relatively cheap to test the glucose level. However, it takes up to at least half day for pregnant women in the clinic, and the other thing is because of the Covid-19 setting, and it is endangered their health.

47:27 in that situation one blood draw becomes valuable. The other thing is, is there any similar product on the market? And from the development aspect we also need to consider.

47:45 And is there any cohort with enough individuals, whether we have a gold standard to compare, especially doing the IVD registration typically, you need to have a gold standard to compare to prove your registration product is valuable.

48:09 One of the key challenges for quantitative IVD registration is reference standard, and the traceability.

48:20 If without good reference standard, it is difficult to define the population, the clinical reference range, which means, if we do not have the benchmark, we will not be able to measure the content.

48:40 And for our manufacture, we reference the GB/T21415/ISO17511 regulations, to establish the procedure for transferable calibration quantity value. A lot of times, especially with the new biomarker, it is not necessary always to have the CRMs which is the reference material, and we must establish in-house the traceability tree to ensure quality control.

49:24 This way we need to make sure our reference standards, and the other standards, are repeatable with other manufacturers.

49:40 For registration IVD product in China here I am showing a Class 2 IVD registration process.

We have registered more than 10 Class II reagent products in 2 years already.

50:00 And this first, we need to have the product design and the pilot manufacturer, and then a clinical trial to evaluate, and certain time if it is a known study and then it could be exempt for the clinical trial. And then the third is technical review for administrative service, and other regulations then followed by the approval and the registration certification for the for marketing.

50:36 And after the market we need to have the GSP and the cold chain logistic to support the shipment and deliverables, and the feedback from the customers.

50:54 Finally, it is the commercialization and with the associate consumables.

51:06 Calibra has developed a roadmap for industrialization, including the non-targeted multi-omics platform for biomarker discovery.

51:19 The targeted multi-omics for confirmation and validation, and also develop the IVD product and IVD registration, and this clinical setting testing methodology, and the reagent and the consumables.

51:40 Thank you very much for your time.

(Q & A – Andrew Percy)

51:47 (Andrew) Excellent presentation, Huafen. Thank you for sharing the results of a few of your study applications and the general roadmap for their clinical translation
.
51:53 (Andrew) There is time now for a few questions, and I see already some from the audience. But before I turn to those, I like to kick this off with one of my questions. In metabolomics, especially data analysis, missing values and sample and replicate data can be found.

52:13 (Andrew) Can you briefly describe your strategy for dealing with missing values? Are these imputed? If so, we are using what type of technique? And what software tool?

52:23 (Huafen) Actually, for a very regulated environment, ff the data is missing, we need to repeat this sample, and to ensure there is no mistake during the sample preparation and also sample collection process. And typically, we will not do the data mining and the find those data and there are some opportunities over there, so that's part reason we need to repeat those samples.

53:00 (Andrew) that makes sense. Thank you for that. Turning to the audience. First question from Liang Zhao. This turns to the profiling that on the proteomics and metabolomics study for Covid-19 patients. He is asking how much difference regarding the protein numbers between serum and urine samples in your covid 19 studies?

53:27 (Huafen) Because during that time the sample collection is really challenged because the hospital is focusing on saving people's lives. And ideally for the proteomics study in the blood sample should be plasma because the hospital is getting used to the serum collection. That is part reason we use serum. For the urine sample because we have a large amount, and we were able to condense for the urine sample, we have better quality data, and more protein was identified, and more metabolites were identified.

54:17 (Andrew) Okay, thank you. The next question also deals with the proteomic and metabolomic profiling study. The question asks what is used as a dataset for machine learning?

54:29 (Huafen) The Dataset we were using is from the non-targeted analysis for both the proteomics and metabolomics

54:44 (Andrew) Okay, thank you. The next question for the patients that had proteins in their urine, did they have kidney disease?

54:50 (Huafen) Because of our severe Covid-19 patient, a lot of them have a chronic disease. The majority are associated with cardiovascular disease, diabetes or kidney damage. So, there are some of them that have tiny damage that's part reason we see more protein than what to be expected.

55:21 (Andrew) Okay, thank you for that. The next question from Greg Everson, who says, "Thank you for the presentation. I am from Hepquant. I have 2 questions: if AUC is what you are going after how you define the acceptable range of metabolite for individual tests?"
The second question. "At the calculation of AUC, do you have a minimum model for data point?

55:47 (Huafen) For the first question, if we are using AUC whether we are using single-point for a single metabolite the point of developing the metabolomics biomarkers is using multiple markers instead of a one individual marker, because a lot of times, for metabolomics screening it is before disease. It is during the progress of disease happening. In that situation, we need to see the changes of the enzyme activity, which is indicated by the upstream, the substrate, and then the downstream the metabolite. We also need to include multiple pathways. For example, if we see the insulin resistance, and our body the lipids and also the branch chain, these pathways are active. In that situation we need to consider if the glucose pathway is down regulated, and if the other password is up regulated, and see the correlation as indication of the disease rather than one single biomarker.

57:28 (Andrew) Lots of considerations. So, we are coming up to the top of the hour and I am cognizant of the time, I want to leave a few minutes at the end for closing remarks.
So, with that in mind, I will close the Q. A. session. Any remaining questions will be addressed through email. I would like to thank you once again all for attending. Huafen thank you for your presentation. I will now turn the podium back to Greg Whitney.

Closing Remarks – Greg Whitney

57:53 (Greg) Thank you, Andrew. Thank you. Huafen very appreciated. Huafen, if I could kindly ask you to stop sharing. I just have a couple of quick slides to close us out here today.

58:06 (Huafen) Sure

58:07 (Greg) And thank you for the wonderful content it is always very interesting to see mass spec complementing and supplementing some older techniques. including the machine learning aspect, obviously the utilization of stable isotopes. It helps us all tremendously, so very appreciated.

58:25 (Huafen) Thank you.

58:26 (Greg) Okay. So just to finish us up, just a couple of quick final statements from my side. Hopefully, we have achieved our goals today, and you have gathered some interesting knowledge and input from Huafen. Also strengthen your understanding of how CIL and our products play a role in these application areas. And as I understand, just from the Q&A in the chat that there are several additional inquiries, so we will endeavor to get those over to Huafen for some responses in short order. And of course, over the last 60 min or so hopefully we have kept you engaged, and you have had some fun learning a little bit more about these techniques.

59:13 We would encourage you for those of you that are attending to come join us at ASMS coming up the week of June 5th, in lovely Minneapolis, Minnesota. Krista, Andrew myself and several of us will be there in attendance at Booth 201. We would love to make some personal introductions and meet those of you that may be attending.

59:35 And finally at the closure of today's session, there will be a quick 3 questions just helping us to gauge the relevance of today's content to your research. And we will also be sending along a follow-up link to the recorded version of today's talk and some additional relevant information.
Of course, we would be delighted to address any product related questions, requests or needs for assistance of any kind from a CIL perspective, we would be delighted to connect with and assist anyone interested. So, thank you again for your attendance as always, we respect your time. Look for that final survey coming in just a moment, and please join us, throughout 2022. We have several additional events planned in this format.

1:00 So thanks again, thanks to Huafen and Andrew, and have a wonderful day.