

**NWX-NIH OD IMOD OFACP**

**Moderator: Francis Collins  
September 17, 2015  
12:00 pm CT**

Coordinator: Stand by. At this time all lines will be on listen-only mode for the duration of today's call. The call is also being recorded. I'd like to introduce (Gretchen Wood).

Francis Collins: Well, this isn't (Gretchen Wood), this is Francis Collins. Good afternoon or good morning, depending on where you are. Everybody thank you for joining this very important phone call, convened by the Advisory Committee to the Director, where we have some substantive issues to discuss.

One on an application - three applications to the HeLa whole genome sequence. And then a discussion about a presentation of a design plan for the Precision Medicine Initiative, a cohort which our very hardworking working group has now put forward for consideration by the overall ACD.

Let me start by doing a roll call of the ACD members so that I know who has joined us. So, please speak up if I call your name. Huda Akil?

Huda Akil: Yes.

Francis Collins: Russ Altman?

Russ Altman: Here.

Francis Collins: Cori Bargmann?

Cori Bargmann: I'm here.

Francis Collins: Mary Sue Coleman? Lisa Cooper? Eric Goosby? Helen Hobbs? Harlan Krumholz?

Harlan Krumholz: Here.

Francis Collins: Cato Laurencin?

Cato Laurencin: Here.

Francis Collins: Rick Lifton?

Rick Lifton: Here.

Francis Collins: Ian Lipkin?

Ian Lipkin: Here.

Francis Collins: Peter MacLeish? Elba Serrano? I think she's going to call in slightly late. We heard that is maybe why she's not currently responding. Moncref Slaoui? Michael Welsh?

Michael Welsh: Here.

Francis Collins: And Chris Wilson. Chris was also going to call in a little later, but is expected to join us.

Well, thank you everybody. A few logistics to mention. This tele-briefing notice was published in the Federal Register on August 20, and amended on September 15. The tele-briefing is open to the public and listen-only lines are available and I hope are being utilized.

Public dial-in numbers with passcode is available on the Web site for the ACD. And also on the site are today's presentations and background material. There will be PowerPoints that were used for the presentation. So if you are listening in and want to go to that site, this would be a good moment to do that. You can just get to it by Goggling, NIH ACD.

This call is being recorded and will be transcribed, so please speak clearly and identify yourself when you make comments. Appreciate the Advisory Committee members doing so.

So those are the logistics. I just wanted to reflect a little bit on some general issues because the ACD has not met as a group since we were all gathered in June, and there's a lot going on. I just thought I'd give you a quick snapshot of some of the things that we here inside the Beltway are tracking rather closely.

And that would include what's going to happen to the budget. And of course this is the time of year where there's generally more uncertainty than we wish there was, and this is no exception.

As you know, we have seen, in both the House and the Senate, going through the regular order process, appropriations subcommittees actually coming up with marks. And both the House and the Senate are proposing significant increases for NIH.

The House at one point, \$1 billion; and the Senate at \$2 billion. And we are very grateful to the broad support we've seen in both houses of Congress and both parties for the importance of medical research, and seeing it reflected in that way.

Certainly if something along those lines were possible it would be a great step forward in what has otherwise been a pretty difficult 12 year slide in NIH's support, accentuated particularly by the sequester.

But what exactly is going to happen is still very much up in the air. There have been some interesting developments in just the last 24 hours. As you may have heard, there have been threats that this might be one of those years where no agreement could be reached, even about a short-term budget resolution.

And of course the government would in that case, on October 1, have to shut down which was very painful experience we had two years ago, and hope very much will not happen again.

The President spoke publicly and quite explicitly yesterday about the dangers of a shutdown. Interestingly, Senator McConnell, the leader of the Senate has said he fully expects to enter into negotiations, and that will both make sure that the government keeps going, and also that the sequester would be lifted.

And of course that would be a wonderful boon, because without lifting of the sequester in case you thought it was over, it's actually the default is to continue it.

If that's not lifted it would be very difficult I think, for the Congress to come up with the kind of positives that I mentioned a minute ago.

The appropriators have said they are working on a short-term continuing resolution to give more time for a longer budget deal. They said they want to be ready. There is, of course, a cloud on this horizon because no one has yet convinced the House conservatives to support such a plan for a short-term continuing resolution unless Planned Parenthood is de-funded, which is something which would probably have very little chance in the Senate. And certainly the President has indicated he would veto.

So there are clouds still very much, although one does always hold up hope that when the rhetoric gets over, some practicality will appear.

So we will see in the course of the next two weeks about what exactly can be done to make sure the government keeps running. If in fact things play out the way that we might hope, we would like to see a short-term CR.

And during that period a Murray Ryan kind of working group as was done two years ago, could be put together with working out the spending levels, not just for a year but for two years, in a fashion that does in fact give relief from the very onerous burdens of the sequester.

But obviously many moving parts before one could say in any confidence, that that will be the outcome. So, stay tuned.

Otherwise, in terms of things that you might want to know about, ACD members, the Common Rule which we've talked about and which has been under intense revision efforts within the administration over the course of some five years, has now had its Notice of Proposed Rulemaking released, and it is out for public comment. And certainly we hope there will be comments on that.

So it is a bit of a length document. It's not one of those quick reads, but it's really important. And of course it will play out in very significant ways for the main topic we're going to come to in a bit about Precision Medicine.

So definitely if you have not heard that that's out for comment, spread the word. I think we have another six weeks or so for comments to come in. So we're not quite up against the deadline.

Woman: Even more than that; 90 days.

Francis Collins: Ninety days and we've only passed probably through 20 of them. And then - sorry, is somebody raising a question? Maybe they were sneezing. Okay, thank you.

Finally, just in terms of items of interest, we were delighted to hear a couple of days ago that the President has now nominated Rob Califf as the new Commissioner of the FDA.

Rob is somebody we all know well. He has been a faithful, wonderful participant in the Precision Medicine Initiative working group. And certainly many of us who have known him for a decade or more have greatly admired his leadership in clinical research. All the work that he did at Duke since coming to FDA as the Deputy.

We've had many opportunities to work with him collaboratively, and have found that to be a very satisfactory and satisfying relationship because of his experience and his wise approach to problems and his collegially.

So we are hoping that the Senate will quickly act to confirm that nomination. And congratulations to Rob who may very well be on the phone. I don't know. He's been a very important part of the working group that's going to report here shortly.

So, I think those were the preliminaries that I wanted to put in front of you. And maybe I'll stop and see if anybody has any questions about those items, or anything else you want to raise before we get into the two main topics.

Hearing none, then I would like us to move to a report from the HeLa Genome Data Access working group. And the ACD members are familiar the way in which we have conducted this oversight of access requests to the HeLa Whole Genome Sequence. And we have three more of these for you to consider. And I'm going to turn this over to Kathy Hudson.

Kathy Hudson: Great, good afternoon. So I am going to be walking through a couple of slides which were sent out to the working group and which are posted on the ACD Web site as well. So I'll try to refer to the slide numbers as we go through this.

So this will be a brief report and then a request for your approval of three data access requests.

So on Slide 2 is just a reminder of the terms of the agreement that we reached with the Lacks family, and which we have asked all researchers, irrespective

of their funding source, to apply for access to the HeLa Whole Genome Sequence that resides in dbGaP. And that they abide by the terms of the Data Use Agreement which are laid out on Slide 2.

Moving to Slide 3, it reminds you of the role of the HeLa Genome Data Access working group which is to evaluate these requests to determine whether or not the requests are consistent with the terms of the Data Use Agreement. And then to report their findings to you and for you to make recommendations to Francis about that.

We have been committed since the beginning in turning around these data access requests quickly, and we have actually been doing a really good job about that. So it's helpful that we have this opportunity today to get your recommendations on the three pending data requests.

Slide 4 is a roster of the HeLa Genome Data Access working group. It is Co-Chaired by Clyde Yancy and myself, Clyde being a former ACD member. And two other ACD members, Russ Altman and Lisa Cooper are members of this group as well.

And then you can see a couple of members of the Lacks family are also important contributors.

The next slide, Slide 5 lists evaluation criteria that we use in evaluating the Data Access Request in order to make sure that they are consistent with the agreement that we have with the family. I won't read through these. This is just as a reminder to you.

And then in Slide 6 are the three kinds - four kinds of findings that the working group takes in response to specific Data Access Requests.

One is that we find the Data Access Request consistent with our agreement with the Lacks family. Second, we can find the request inconsistent. Third we can make a conditional finding which is basically that there's some missing information that can be obtained by a staff. And if the requestor satisfies that, we can go ahead and make that approval. And then last is pending, which requires a reevaluation by subcommittee.

So on Slide 7 is the status of the requests that we have evaluated up until now. We have evaluated 43 requests which is actually a robust number of people who have been seeking access to ad using the sequences.

You can see there the numbers that have been disapproved, either because they didn't respond to our request for more information. One was disapproved - recommended for disapproval by you, and subsequently disapproved by Dr. Collins because of the failure of the group to want to share the results of their research.

On Slide 8 are the three current proposed projects. The Data Access Requests themselves were put up into the SharePoint site that you had access to. These were reviewed by the HeLa working group.

Usually in the discussions we rely on the Genomic members of our group - Rick Myers, Bob Nussbaum, and Russ Altman to provide some lay explanations if we need that.

So the three proposals are on one Human Spliceosome Structure and Function; one on Copy Number and HeLa Cells Copy Number; and a third using Hi-C to look at Genome Architecture. So those are the requests.

The working group found that all three of these requests are consistent with the Data Use Agreement. And so we would put before you and ask for a motion or discussion to make a recommendation that Francis provide access to the HeLa Genome Sequence for these three studies.

Francis Collins: Great. So, ACD members, looking for questions that you might want to pose about any of these. Hearing none might I then ask for a motion.

Cori Bargmann: Moved. Move to approve the three.

Harlan Krumholz: Second.

Francis Collins: And I'm sorry, who was the move and the second?

Cori Bargmann: Cori, moved to approve.

Harlan Krumholz: And Harlan, second.

Francis Collins: Great. And let me just check and see whether anybody has joined who's an ACD member since I last called the roll so that I know who we have.

Did Elba Serrano come on yet or Chris Wilson? Fine, we still have plenty of people.

So I'm going to just ask for a voice vote and then we'll see where we go from there. So, all in favor of the motion please say aye.

Woman: Aye.

Man: Aye.

Francis Collins: Are there any opposed? Are there any abstentions? Great. Then I think by this mechanism we know who voted. And thank you all very much.

Kathy Hudson: Thank you.

Francis Collins: And thanks Clyde and Kathy for continuing to lead this effort. And I'm sure we will have other proposals to bring forward the next time we are together.

Okay, well we're doing pretty well here. Let us know move on to the other main topic which is the Precision Medicine Report from the working group. This obviously has been an incredibly intense effort over the course of the last five or six months.

I particularly want to thank Rick Lifton, Bray Patrick-Lake, and Kathy Hudson as our working group Co-Chairs for having led this effort, but many others as well.

And among the people that I particularly want to single out, I would want to thank Dr. Gwynne Jenkins who has served as - I don't know what your title is; the person who just does everything for this particular working group effort.

Around the table with me here at NIH are in fact a couple of other folks who are part of the working group. Josh Denny who you'll be hearing from shortly because he's going to do part of the presentation, and Eric Dishman who happens to be here because he was a hearing yesterday in front of the Senate talking about access to usable electronic health records and all of that. So, appreciate their being able to be here with us in person.

So I think I should at this point turn this over to the Co-Chairs with many thanks for all the hard work you have done. And let's get into the presentation and I'm sure the ACD members will have questions and want to discuss this.

So this is a rather seminal important milestone; an inflection point for the President's proposal to do something really historic. So please, let's hear what they've come forward with.

Kathy Hudson: Great, so I'm going to kick this off and then we're going to do a little bit of a tag team among the Co-Chairs and with Josh Denny. I'm going to be working slides, and we'll try to refer to slide number.

The slides are up on the ACD Web site, as is our report. And of course you as the ACD members have had access to the report for a week or so now. And I hope you've had an opportunity to review it.

So on Slide 2 is just a reflection on how this effort got underway. In the State of the Union Address on January 20, the President announced that he was launching the Precision Medicine Initiative.

And ten days later in the gathering with many patients and scientists and others at the White House, he talked more about what his hopes and aspirations were for this Precision Medicine Initiative.

On Page 3 is the sort of the mission of the Precision Medicine Initiative that's been put together by the White House. It is not just about the National Institutes of Health, although we have a substantial portion of the proposal's that are in place. But it's really to empower patients, researchers, and providers to work together to develop individualized treatment.

On Slide 4 is the budget request that the President put forward for FY'16. And we, despite the sort of gloomy situation on The Hill right now, we remain optimistic that there will in fact be a budget and that it will include these sums for the Precision Medicine Initiative.

It is exciting that both the House and the Senate Appropriations Committee have included the \$200 million in their Committee bills. They're marked for FY'16. And so we are optimistic about that. It feels like there is a strong bipartisan support on The Hill supporting this initiative.

So the Precision Medicine Initiative research components are twofold. One is the research cohort which we will talk about today. And that has to do with all diseases.

The second is focused specifically on cancer and really on cancer clinical trials and related activities. And we're not going to be talking about that today.

The FDA component is, in order to enable Rob and his colleagues at the FDA to develop nimble, flexible mechanisms to be able to review genomic technologies and specifically genomic technologies around diagnostics.

The Office of the National Coordinator or ONC has been proposed to receive \$5 million. That's really related to interoperability of electronic health records, and blue button functionalities to be able to allow individual patients to easily receive their own health information.

And we'll come back to that because it's really a critical piece of what we need to have happen in order to make the Precision Medicine Initiative truly work.

On Slide 5 is the general charge that was provided by Francis to this working group when we were put together back in March. And that was to develop a vision or a blueprint for this cohort program that will be a national, longitudinal research cohort of a million or more volunteers.

And some of the questions that we have been wrestling with over the last month together as we've gotten to know each other quite well is, do we want to leverage existing cohorts or start from scratch, or some combination of those?

How can we effectively capture the rich diversity in the United States? What kinds of data should we include at the beginning and as we go forward? And then what kinds of policies need to be in place for this to be maximally beneficial to our researchers and research participants and biomedical research.

So the next slide; Slide 6 shows the working group members. As Francis said, Bray, Rick, and I have had the pleasure of leading this group together. We have had a really interesting, diverse group of people, many who had not had the opportunity to work together before, coming from many different areas, disciplines; etcetera.

And so these folks have really given up their spring and summer to dig in on this project together. And it has been quite an adventure and really an honor for me to be a part of.

On Slide 7 is a summary of some of the inputs and mechanisms of getting inputs along the way. We really wanted to make sure that we were tapping into the broad range of experience and expertise on a number of topics. So we

had a series of four public workshops. They were Webcast so that anyone could tune in.

Those Webcasts are archived on the NIH Web pages and so they are available for folks to see. Some of them were WebEx, so we were taking comments from people who were participating remotely.

All of these had a remarkable Twitter presence. And I think Bray won the highest number of tweets during several of these. We were trending across the United States during some of these workshops at #PMInetwork if you're paying attention to Twitter.

So those are the four areas that we had as workshop topics listed there. Really appreciated both Vanderbilt and Intel hosting us for two of those workshops.

We also had requests for information that went out, one about for cohort capabilities that are out there and import omens in building a cohort. That was very useful. A large number of folks who have cohorts responded to those and provided some helpful insight.

And then a second request for information about how to reach community groups, engage communities, and particularly tap into underserved and under-represented communities.

We also had the opportunity with help from the foundation for the NIH to put into the field and get data back in really short order for some immediate gratification, a survey of public perceptions about Precision Medicine. And I'll refer back a little to some of those statistics further along.

And then lastly we had as an input to our proceedings and deliberations the work product of a group that the White House put together to really focus on the issue of privacy and trust. And that was something that the President mentioned early on.

Of course when his was launched that was before one of the largest privacy breaches in government data, and so this issue has only become more important over the course of our deliberations.

And those White House privacy and trust principles have been out for public comment, and we certainly took them to heart and they are referenced throughout our report.

So with that as sort of the inputs, I'm going to now turn this over to Rick who's going to talk about the next chunk of slides.

Rick Lifton: Great, thank you Kathy. At the outset I'd like to add my thanks to first my Co-Chairs, then the members of the working group and the remarkable staff from NIH who really devoted incredible time and effort and wisdom and insight to bring this to its current stage.

In addition, at these workshops you know, literally people were invited on the fly with almost no time. And virtually every invitation was accepted. People really dropped everything to participate. And people came from all across the country to participate in these workshops, as well as the people who signed up to attend the workshops.

It's really been a remarkable experience over the last five months; a very short timeline.

And among the fundamental questions that we were posed with are, why should we do this and what are our expectations coming out of it? And there was unanimous enthusiasm for this project, largely based upon the recognitions that we still are remarkably naïve about the passage genesis of most of the common diseases in our population.

We lack fundamental knowledge of the causes, both genetic, environmental, and demographic factors that contribute to disease. And in particular we are very limited in our ability to prospectively identify people who are at highest risk of developing a wide range of diseases.

And if we look at why we have been so successful in reducing the prevalence and impact of coronary heart disease on the general population over the last 50 years, it was largely attributable to the prospective identification of smoking, high LDL cholesterol, and high blood pressure as major risk factors for this disease that could be mitigated.

And this of course came back to the establishment of a prospective cohort where you could identify factors before people came down with a disease that might be mitigated by therapeutic intervention.

And we haven't really been able to replicate that success in the large number of other diseases, in substantial part because they're so much less prevalent so, you would need individual cohorts that were tailored to study the incidence of individual diseases.

So because of that the idea has been that we might be able to do this in a single, very large cohort of people from across the United States and collectively study the factors that ultimately lead to the development of a wide

variety of the most common disorders in the population in a single large cohort. And this is a major motivating factor for this project.

As well, we recognize that when new therapeutics are developed, they're developed to be of net benefit in the populations that have been tested for them. Recognizing that many therapeutics will have great benefit in some, no benefit in others, and may actually have adverse effects in yet other individuals in the population.

And so we have been relatively in the dark about the factors that contribute to individual variation in response to therapy. And with a sufficiently large cohort where people have prevalent and incident disease that ultimately gets treated, we ought to be able to learn a lot about the determinants of individual variation and response to therapy.

A third point is that most studies in the United States have focused on small segments of the population that may not be generalizable to the entire population. And certainly the Caucasian population has predominated - the White European population has predominated in clinical trials in this country.

And that leaves open the question of whether the same therapeutics and preventive strategies apply to other populations in the United States.

Similarly, socioeconomic factors play a very large role in disease outcomes, and we have very little experience in sampling across socioeconomic categories.

And it was pointed out in one of the workshops that a major risk factor identifiable in the country for morbidity and mortality across a wide range of diseases is simply zip code, because that's strongly correlated with

socioeconomic outcome and is an area that we think is important to investigate.

So the goal of this cohort then would be if we sampled a very large cohort of people in the United States, that we would be able to accelerate research across many areas of health and disease.

And then finally, it's recognized that in order to be successful in this effort, we would like to have longitudinal participation of the participants in the cohort, and would very much like the ability, when new findings are made that may be relevant for understanding disease and health biology, to be able to go back to those individuals and do further studies.

Either to link genetic or environmental factors to pathophysiology that contributes to health and disease, or alternatively, to think about clinical trials that might be relevant to particular subsets of individuals with particular biomarkers that appear to be driving their disease.

So on Slide 9, the question then is, why should we do this now as opposed to any other time in our history?

And we recognize that there are convergent factors from diverse areas of technology and inquiry that really make this seem to be an opportune time to do this.

And so as shown on this slide, when we were sequencing the human genome ten years ago, it's estimated that it cost many millions of dollars to sequence the human genome. And now most places that are doing this at any capacity would say, we can sequence a human genome for less than \$2000.

Similarly, the time has gone from a couple of years to generate the sequences required, to again, a day is plenty of time to produce the sequence for a human genome today.

Similarly, it's critical that a project of this scale would be able to collect data in a very automated, passive mode, preferably. And part of that is first, the number of smartphones in the United States, now nearly 60% of the population have smartphones.

And moreover, that electronic health records, which were really just coming into wide use ten years ago, now are used by about 90% of healthcare providers; by more than 90% of healthcare providers, providing tremendous opportunity to collect data in a form that in principle can be computed upon.

And then lastly the computing power itself has drastically changed about 16 fold higher raw computing power.

But again, as we heard in one of our workshops from Russ Altman, the progress in deep learning where orthogonal data sets can be combined and computed upon. And important insights that would have been very difficult to produce by other means are really emerging with quite remarkable results.

So, for all of these reasons it would have been almost inconceivable to imagine putting together a cohort of a million people with their complete medical records ten years ago, and combine that with genomic and broad environmental data. It seems very feasible today.

So the scientific opportunities, we believe that this cohort has the potential to discover -- now on Slide 10 -- to discover new biomarkers predictive of future disease risk which will provide opportunities for new approaches to not just

disease treatment, but hopefully disease prevention. Discover the determinants of individual variation in response to therapeutic; enable us to really in the real world setting, determine quantitatively the risk estimates in the population that integrate genetic environmental and gene environmental interaction terms.

The development of mobile health technologies we also think has tremendous opportunities for application to this study. It became clear from our workshop that simple measures of activity can provide incredible insight into the development and progress of a variety of neurodegenerative disorders.

And simple measures of activity and sleep for example, we thought would have the potential to provide fundamental new insight into a wide variety of factors that contribute to health and disease in the human population.

We also think that if we have this large data set with genome sequence, that there will be a very large number of heterozygous, loss of function mutations in each of the roughly 20,000 genes in the human genome. And that there will be enough power to assess what the clinical impact of those variants might be.

And it will be able to identify new classifications and relationships among diseases because we're starting with an unbiased sample in which we won't have selected patients for particular traits at the outset.

Lastly, we'll be able to do targeted clinical trials in patients with very rich clinical data, which has not typically been done previously.

And then truly lastly, a major goal of this initiative we feel should be to make this data broadly accessible to the investigator community so that these very large data sets can be computed upon by a wide range of investigators asking

different questions, using different methods, and this could provide an incredible resource for advancing the health of the nation.

In terms of nuts and bolts, if we just think about the estimated disease incidences and prevalence on Slide 11, in a cohort of a million people -- I won't take you through all of the details except to say that we expect to have a quite rich set of individuals, both with prevalent disease at inception and with incident disease at five to ten years out.

And in the full document we provided power calculations indicating the power to detect significant effects of genetic environmental and gene by environment interaction.

So then on Slide 12, to assemble the cohort, our fundamental notion was that the individual should be enrolled with, be re-contactable if there are interesting findings. That we want to go back to them that we should be able to do that.

That we should be collecting complete electronic health record data, biospecimen surveys, and a baseline exam. This will be a longitudinal cohort with continuing interaction, and additional samples will likely be collected as the cohort moves along.

We envision two methods of recruitment, and this was something that we spent a lot of time on.

We think it has the potential to be a galvanizing effort for a national research purpose if in principal, anyone in the United States could sign up for the cohort. And we think that's a very attractive idea. However, it poses a number of questions about how scalable is that? The ability to re-contact these

individuals and follow them longitudinally may be more challenge than individuals who are recruited from healthcare provider organizations.

And as the ACD well knows, there are a number of healthcare provider organizations that collect virtually comprehensive healthcare data because all of the healthcare that individuals receive are through these single entities such as Kaiser Permanente.

And there's this strong feeling among the working group that groups that collect large healthcare records and with longitudinal follow-up, should comprise a very healthy chunk of the cohort as it is developed.

But key to this of course, given our other goals is that we need to make sure that the ascertainment and recruitment of participants achieves the diversity that we're seeking in the cohort.

Then on Slide 13, we want to make sure that we broadly are representing the diversity of the United States in the cohort. That traditionally under-represented groups are well represented; that we have all states of health and diseases; that we broadly sample across the United States, across all life stages.

And this latter term requires special policy considerations regarding enrolling children; individuals who are or become decisionally impaired, and participants who may become incarcerated during the course of the project.

So these are a number of the considerations that went into our deliberations. And I now will turn it over to Bray to discuss a number of important aspects of participant engagement. Bray.

Rick Lifton: We can't hear you. Hello?

Man: We can hear you Rick, but not Bray.

Coordinator: If Bray is on line, have them press star zero. This is the operator. One moment. Bray's line is open.

Rick Lifton: Bray, are you there?

Coordinator: Bray, can you press star 0 so I can open your line up? This is the operator.

Man: Rick, it sounds like they lost the connection at the NIH. This is (Sechin). So if you want to continue the slides, Kathy was saying to continue until they are able to reconnect.

Rick Lifton: Okay. So on Slide 14 I want to - I'll just continue until Bray is able to connect again. President Obama in his talk on January 30 said, "I'm proud we have so many patient rights advocates. They're not going to be on the sidelines. It's not going to be an afterthought. They'll help us design the initiative from the ground up, making sure that we harness new technologies and opportunities in a responsible way."

And this really set the stage for much of our discussion, really informing us about the importance of participants being empowered to help drive the research agenda and play a major role in determining how this is handled.

On Slide 15 is the results of a survey from FNIH of public opinion on a large cohort study. And there was broad consensus that this is a good idea. It was interesting that 82% of respondents were interested in receiving the results of the study. And 62% want to help advance health research. And there was

broad support for the idea that participants should be partners with the researchers.

On Slide 16 is the - has Bray connected? It doesn't sound like it so, I'll continue on Slide 16 on the focus on engagement.

And so there's very strong consensus to have an interactive and proactive participant model. That participants in the cohort would be well represented in the governance, design, conduct, dissemination, and evaluation.

Critical to that is having a strong foundation in trust. Important to have participant engagement and communication activities are centrally coordinated so that individual sites and individual investigators are not going back to participants, other than through the Central Coordinating Center.

And then importantly, that the consent should be, with the Precision Medicine Initiative so that although we will be recruiting a large number of participants from healthcare provider organizations, those organizations will be involved in the recruitment. But the recruitment - there will be new consent for each participant that will be part of the Precision Medicine Initiative cohort program.

Francis Collins: Hey Rick, we're back.

Rick Lifton: Terrific. I am so relieved.

Francis Collins: Thank you for keeping the ball rolling while we dealt with a complicated and difficult conference setup. So I don't know, where are you?

Rick Lifton: So we are just at the end of Slide 16 where we have gotten through the consent issues and were just saying that there was going to be a single IRB for the cohort. And I'll turn it gladly over to Bray.

Francis Collins: All right.

Bray Patrick-Lake: Hi, Bray Patrick-Lake. So it sounds like we are at the end of the slide where we're talking now about the return of results and access to data.

So the working group felt very strongly that information sharing with participants was an incredibly important aspect of making this a successful initiative. Aggregate results should be made available to all participants.

And the recognizing that there's no, one size fits all model for participants and their preferences, we believe that individuals, to be able to set their preferences for the return of information.

And with that I believe I am now passing it on to...

((Crosstalk))

Francis Collins: Bray, you were just great.

((Crosstalk))

Bray Patrick-Lake: Thank you folks.

Josh Denny: Great, this is Josh Denny. I'm going to talk - so I'm now on Slide 17. And this details how participants and data and biospecimens would fit into the cohort.

And we have two - or I guess three columns here. And on the left we have in blue, that people who are direct volunteers that are not coming in through a healthcare provider organization or HPO.

So far right we have on green we have the people coming from HPOs. And in the middle we have the data that would be - and specimens contained with PMI.

And so for - and then we have arrows to whether the data is coming from the individual or whether it's coming from the healthcare organization or a physician.

And so you can see some important things here. You can see first the key elements of data that we look to get. So report measures, mobile health data, consent of course, EHR data.

Now notice EHR data for the direct volunteers comes from a different area than from the healthcare provider organization. And this point we're saying the direct volunteers will be in charge of and donating their own healthcare EHR data. And others could come from insurance companies or Center for Medicare and Medicaid Services.

The technologies such as we've gotten - and we actually spent a lot of time in the report discussing how that data would be shared and what evolutions would be required to existing technologies specified in meaningful use too, to really make it effective.

In addition, individuals from both sides would provide baseline exams and biological samples. And those would go into a biobank. And we have said, for PMIs, those really should be new specimens that are collected specifically for

PMI. If there is an existing HPO related biobank that would just be independent of PMI. Next slide.

So this slide, 18 details our information out and interactions with researchers with the data. And then how data would also flow back to the volunteers in PMI.

And so in this case we have our biospecimens at the bottom and our data Cloud on top. And we would have researchers that would post questions that may use biospecimens and data in the Cloud. And then as they generate new data and information from that data, that would be redeposited into the central PMI data source.

Francis Collins: All of them using their Erlenmeyer flasks of course.

Josh Denny: All of them using their Erlenmeyer flasks.

Kathy Hudson: And wearing goggles.

Josh Denny: And wearing goggles, yes. We - yes, nothing else to say on that. So you can see - and the data also goes back to the participants.

First off they already have access to their clinical data. And so one of the key things we heard in the FNIH Study was that participants want access to their clinical data. And PMI we believe can be a catalyst to aggregate and present back to individuals, their clinical data from HPOs.

And those donating their own data via blue button and derivative technologies, access that way as well.

Of course everyone should have access to aggregated results and studies that are done at PMI, and have ongoing study updates. And we mentioned the individual health data which could extend beyond just EHR data, but also insurance data and other forms.

And then in general, individual data can be generated from a variety of technologies. And we want to preserve the individuals really to access information. But you know, not all information is of the same grade and quality. Not all testing will be done or could be done in CLIA certified labs.

And so the working work discussed quite a bit, what return of results in information looks like in a variety of contexts. Those results that are done in CLIA labs that are considered actionable are ones that we believe should be returned to participants. And they should be - or I should say, they should be able to receive that information if they so desire, and set preferences to receive that information if they wish to do so.

And then other classes of data that are maybe more research that don't have as much validity behind them, may be need to be made available to the participant when they acknowledge that maybe this isn't done in a CLIA lab and can't even verify for sure that the data contained, is from me. But you know with those provisions they could still be able to access their data.

And additionally we recommended having a subcommittee that included participant involvement be formed to investigate, oversee; to be involved with these decisions on an ongoing basis. Because we foresee it being an area of active discussion throughout the early parts of the project, and one that will be dealt with in many perspectives.

Now Slide 19, we're talking about the possible data sources for part of the cohort. And one principle was that we wanted to think large in terms of the kinds of data that we might want to have in the cohort over time, and not just what is available today.

So this categorizes a number of different tapes that we foresee being part, starting with self-report measures, clinical data for EHRs, and other healthcare claims data, come in both structured and unstructured forms. We want to preserve the ability to access both of those.

Included in unstructured clinical data, of course the narrative documents will also form data - EEG data, EKG data, imaging data that represent large data sets that could be aggregated and new ways to explore health.

We talked about biospecimens, mobile health and sensor data. We included a lot of things that could be research grade, as well as commercially available, that we - many of us may be wearing right now or carrying in our pocket; a smartphone.

And we want to leverage both of those, from the outset taking available technologies such as smartphones and commercially available technologies would be early forays into these technologies and use.

I mentioned healthcare claims data earlier. It is an important aspect that could be gathered both nationally from an insurer such as CMS or Blue Cross, and others. And then allow a way to follow individuals if they attend or see other healthcare systems.

(Unintelligible) on environment data is an important component that could be linked in with your location information. And then of course other data such

as social networking are things that we wanted to talk about and consider as well. Next slide.

We talked a lot about the initial core data sets. So, where do we want to start from time zero to collect and make available for analysis within the cohort?

We believe that a core data set should be centrally located and stored within a coordinating center. This would facilitate access to the data, it's permanent, and on the speed at which the data could be queried.

We felt it was important to align this core data set with other data sets where possible. Looking at national and international cohorts that linked to biospecimens gives us some guidance as to what kinds of variables to include. And would allow for this set to be able to be combined for Meta analyses in other studies more easily.

And then of course we should allow the existing data standards and data models; common data models where possible, and a number of standards out there that we discussed in the report. I won't go into detail now.

So within these are specific things we want to (unintelligible) include a set of self-report measures that would include things such as diet, substance use, and (unintelligible) and disease and symptoms. A baseline health exam for all individuals, whether they be a direct participant or one who comes in through a healthcare provider organization. And this would include just a set exam, bio; etcetera.

We wanted to capture all the structured clinical data from healthcare provider organizations and list out explicitly what we expect them to be able to send for an individual to be a participant.

We of course want to have some biospecimens, and at the beginning be suggested that synched with blood, but also recommend a subcommittee to look at other types of biospecimens that you want to collect.

And then finally, in-health data. We felt like within the likely, sizeable subpart of the cohort, we could collect available data from smartphones or other fitness technologies if they made a purchase, to give us early entry to using the data from (unintelligible) to the cohort.

Next slide; Slide 21, we're looking at the data flow now between the Coordinating Center of participating sites and researchers who want to ask questions now against the (unintelligible), you know, researchers with the (intelligible), asking data - asking questions of the data again. And you could also be a computer scientist. These could be high school students. These could be a variety of individuals. We want to open these data up in a large form to many individuals to ask questions of.

And so, they would pose questions and many of those questions probably, we believe, could be handled by data with - that is housed at the coordinating center -- that we would consider core data that has been curretted. And over time, the amount of data that is at the coordinating center and the amount of data curretted would grow and as individuals build algorithms to identify diseases of interest, define better medication (unintelligible) and things such as that, that data becomes things that could be easily queried as well.

Some questions, though, are going to require more detailed interaction with either participants directly, or with health care provider organizations of (unintelligible). So these would be things that require looking back at imaging

studies or narrative reports or really detailed (unintelligible) type questions that may come from particular genetic studies.

Slide 22 - on bio specimen collections, we felt like PMI - that PMI should collect new bio specimens. The collection of bio specimens should anticipate what future uses may be driven by scientific (unintelligible). These should be collected from everyone initially and then in subsequent intervals as determined by youth cases and then maybe the protocol figured out for how to collect those. We felt that the cohort should start with blood, but consider a number of other types of bio substance. We felt establishing a central PMI Biobank from day one is important and building the infrastructure to house those samples, and it's important that these samples be collected in a CLIA compliance fashion so that we can maintain a chain of custody of the samples and the testing where possible to enable us to determine most individuals in a way that we can know that this is their sample when it's possible to do so.

And with that, I will hand back off to Kathy.

(Kathy): Great. Thanks, (Josh). So I'm now on slide 23 and one of the conversations that we had throughout the workshops and in the working group deliberations was about what kinds of policy challenges or policy opportunities do we need to think about and what kinds of policies do we need to have put in place in order to make sure that the precision medicine cohort program is maximally successful. And so there is a section in the report -- section seven -- specifically addresses a number of policy issues, and of course many of these policy issues are not within the power of NIH alone to be able to solve or resolve. And yet we think we can work with our sister agencies and with folks on the Hill in order to move some of these policy issues forward.

Among the issues that we talked about included the need to have a single IRB responsible for the precision medicine initiative and that that single IRB include a very healthy representation from participants, unlike traditional IRBs that need a single lay representative or a single community representative. We would like to have a really robust representation of folks on this IRB.

We also talked about privacy and security policies. Both the need to be able to have well defined, data security standards and to be able to protect against unintended data release, to have notification if there is a data release that happens, and then to put in place penalties for unauthorized uses of the data, including attempts to re-identify study participants.

The precision medicine cohort program is envisioned to broadly embrace data sharing goals -- and so that the data would be widely available as was discussed. We talked about supporting the broad consent proposal that has been put forward now in the common roles notice of proposed rulemaking that Francis mentioned earlier. The administration is strongly pushing that the common roles are finalized quickly and having it in place so that this initiative can use broad consent for secondary use of specimens and data is one important catalyzing feature.

And then we talked a lot about data policies in terms of interoperability of electronic health records, blue button functionality so that individual participants can get their data and share it with PMI, and also to make sure that people have the right of access to their own information, and that requires some clarification of policies around the clinical laboratory improvement amendment and HIPAA.

And then there were also issues around certain populations that need some special consideration in order to ensure that we can effectively give them the opportunity to participate in the cohort and be able to retain them over the course of the longitudinal study.

So turning then to slide 24, we talked about governance for the PMI cohort program. Specifically, we spent time talking about the attributes that we would like to see in a PMI cohort program director - really want to see this person have scientific stature, real leadership and authority accountable to the NIH director, connected tightly to all of the institutes across NIH because this is truly a trans-NIH endeavor. And that that person would lead a steering committee, which would have an executive committee that would be small and nimble, and that the steering committee would represent all of the moving parts of this cohort that (Josh) and (Bray) (Rick) described.

There's also recommended that there be an independent advisory council, which really is our board, which is really analogous to the role that council served for institutes and centers. And this model - I think we're really looking at something that's a cousin to the multi-council working group, which has been used both for the brain initiative and for the BB2K initiative. These are programs that are important to many institutes but need their own dedicated council to think about funding plans, to think about planning to watch over and nurture these programs in a concerted way. But also have a degree of independence to provide advice and oversight. And so that is recommended as well.

There were a number of issues that we raised during the course of this presentation and that are described in the report where we think additional concerted attention really needs - and focus is needed. And those, including

which are in results, data bio thinking, access and security really need individual subcommittees as a part of this overall governance structure.

And then importantly, because we are not the only part of the overall precision medicine initiative, we need to maintain what has been really effective coordination with other agencies and divisions within the Department of Health and Human Services, and other departments across government, and of course with the White House. And so that's an important part of the governance as we move forward.

So, moving to slide 25, we now are at a point where we need to consider this report, and then move forward. And we need to think about being able to act quickly to bring on a director to staff this program up, to build the infrastructure, to support enrollment, and to see this get launched officially in FY16, should you embrace this report and recommend it to Francis.

So before I turn it back over to you, Francis, I just wanted to send on slide 26, with a huge thank you to all my new friends and colleagues on the working group, who have been fantastic. Also, to all the people who participated in our workshops, provided comments, tweeted along with us. To the NIH folks who helped plan and implement the workshop. To Vanderbilt and Intel, and especially to (Gwen Jenkins) and her staff. So with that, Francis, I turn it back to you.

Francis Collins: Thanks, (Kathy). Thanks, (Rick). Thanks (Bray). Apologies about that little phone glitch that kept (Bray) from being able to walk through her part, but thank you, (Rick), for stepping in and keeping everything rolling. I know we do have working group members who have dialed in. As I understand it, (Esteban Berchard), (Tony Coles), (Subshah Khaterpaul), (Shiriki Kumanyika), (Pearl O'Rourke), (Richard Platt), (Jay Shendure), (Sue Siegel)

and (Rob Califf) are on the phone, and as I mentioned earlier, (Eric Dishman) is here in the room with us, as is (Josh Dennehy) -- you've been hearing from. And I believe these folks are actually in situations where they can - at least most of them - comment if a issue comes up that they want to weigh in on from responding to what the ACD questions might be. And I'm told there are 450 people listening to this phone conference, so thank you all of you who have dialed in to listen to the discussion of this very important presentation.

So I guess I'd like now to throw this open to the ACD members to ask questions, make comments, whatever your pleasure is. Please identify yourself so people know who is speaking. The floor is open.

(Russ): This is (Russ). I have a question.

Francis Collins: Go ahead, (Russ).

(Russ): Okay, three issues I'll try to throw very quickly. The first is a clarification for (Kathy). I assume that the director you're talking about is an inter NIH person and if this wouldn't be sent outside of NIH - but that's a question.

(Kathy): You want to go ahead and do all three of your questions real quick?

(Russ): Okay. Second question is I'm wondering in particular about the million vet project and if that -- how that played into the consideration. I know it's an existing cohort and it sounds like you want a new cohort. And I just want to make sure I understand that the implications that existing cohorts should proceed, do what they do, but they won't necessarily be folded into this cohort. And then the third issue is - if patients have access to all of their data - this is actually a profound one, I think. If they have access to all of their data, either through Clear Labs or the research data appropriately decorated as

(Josh) described with warnings about its research quality, can - will investigators be able to make direct requests for that data to the participants in order to kind of do citizen science and not through the PMI, go get the data directly from the participants if they decide that the request is reasonable and something that they wanted to participate in? I find that both exciting and it gets my heart rate and blood pressure up as well.

Francis Collins: Ours too. Okay.

(Kathy): Okay, so let me take one and two and then maybe turn to (Rick) or (Josh) or another working group member, (Eric), to talk about three and citizen science.

So, we had a pretty robust conversation about whether or not the PMI cohort program director should be a fed -- federal employee -- or not. I think what we - we wanted the person to be nimble and innovative and have a lot of authority and be able to make decisions quickly and be able to evolve the program as the science changes and as the cohort evolves. And sometimes nimble and innovative are not synonymous with a federal employee.

Man 1: Oh, really? Just saying it how it is.

(Kathy): So, but we did end up realizing that this person has to be able to have direct hands on control to what are the planned funding opportunities that are going to be coming out. What's the appropriate mix of how those awards and opportunities are put into place, and make the final decisions on funds that go out of the NIH.

So we did end up feeling like it would probably - this was a federal position and one that needs to have great stature and authority, and maybe a little different than your run of the mill federal employee.

The second question about MVP, you're right. The Million Veterans Program is a program that started enrolling about four years ago. (Mike Gaziano), who really runs that program, I think is on the call now and has been a member of our working group and a very productive and helpful member of the working group.

So they have 420,000 who have signed up to participate in that program. They have a couple hundred thousand genotypes now. We have been working actively with them to figure out how these two programs will be coordinated, and are excited about the opportunities there. We certainly have learned a lot and will continue to learn a lot because they're out ahead of us. If we can figure out interesting ways to stitch these together that will certainly add further power to the overall enterprise of precision medicine. And who wants to take this interesting, provocative, almost profound question?

Man 2: Well I think maybe (unintelligible). (Rick), go ahead.

(Rick): Sure, sure. Just to follow on to (Kathy's) response, you know, thinking about the directorship, the UK Biobank project has been led by (Rory Collins), who I think is a person of the kind of stature you would ideally want to see in a director for this project - who, you know, has a very broad grasp of the science in many dimensions. And I think would be the ideal kind of person to think of as a leader of this project.

Thinking about your question about getting data direct from participants, seems to me that it's a feasible path if the individual participant, for example, posted the availability of their personal data somewhere and then people could go after it individually. Otherwise, individual investigators would not have the ability to identify individual participants in the cohort.

And, so I think if any individual in the cohort wanted to do that - it's not obvious to me that we would want to preclude that.

Man 1: I think that's a good summary. The individual does have the ability to get their data then what they do subsequent to it through that process of getting it is kind of, I guess, their decision. However, people can - the whole idea of having it - a central entry point into PMI - I think would be able to facilitate all sorts of different kinds of research such that you would presume it would be far easier to do almost any kind of research by coming through PMI to access individuals. And certainly if you wanted to identify individuals -- the only way to know who's actually in PMI -- unless people self-identify -- would be the coming through the core itself.

(Sudha): Hi, this is (Sudha Akeel) and actually in relation to this, I wanted to ask about family data, because this is where I could see an advantage. If you could get a family to share your data in a way that can't be identified in the identified data, if you will, that could be - so this is my concern is also my concern about the children. So if there is, say, a private mutation in a family and somebody was interested in following that through, I don't know whether it's more available if you go directly to the members of that family as opposed to trying to dig that out in a de-identified set of samples.

Man 2: I'll respond to that. So I think that's a really important question. In terms ascertaining and recruiting families, that certainly is one context in which we can imagine enrolling children, and that might go particularly well in HPOs where all members of a family are part of the same HPO and covered, and all of their data is collected in a central fashion. I - and that's one way that you can imagine achieving both enrollment of children and collection of family data.

The interesting question of if you identify an individual with a particularly interesting or provocative mutation, could you potentially follow that out into the kindred - these are, I think, the kinds of questions that we would be open to the idea that this would be part of the re-contact procedure that we've discussed - that you - we could go back with a proposal, for example, through the central coordinating office, to the individual and ask them would you be open to a -- an extended study of yourself or of your family. And provided that they said yes, that could then be pursued by -- either through the central office or through individuals with the permission of the PMI, with the understanding that all of the data would be returned through the PMI.

(Ian Lipkin): This is (Ian Lipkin). I had a couple of general comments and then some that are more discrete - I'll hold for the time being. First (unintelligible)

Francis Collins: Okay, please proceed.

(Ian Lipkin): ...extraordinary report and I'm very impressed. My three general comments are - first is I'm very concerned about the impact of data that may not be accurate, in guiding decision making. The patients and clinicians can make inaccurate estimates of disease risk and they actually choose courses of treatment that may not be appropriate. I don't have a solution for this, but I think it's something that's going to merit some further thought.

Second point is as we get more information and some genetic information is obtained, maybe we need something equivalent to a data safety monitoring board so that we can inform people of risks that we learn as they're in the cohort and we get that sort of information.

And the third point has to do with intellectual property. And I didn't see that either in the slides or in the report. Is this something that you've discussed?

(Kathy): So - this is (Kathy). Let me start and then ask others to chime in as well. In terms of monitoring return of information and monitoring whether or not there's information that needs to be returned, or information that has been returned that may have implications for participants, a DSMB is an interesting model. What we had proposed here was that there be subcommittee of the governance structure that would be specifically focused on that issue. So I think that's where we envision that issue being sort of addressed.

The other questions about concern of inaccurate data being returned to people and somehow affecting care. (Josh), you want to...?

(Josh): Sure, I'll take first stab. The - so a couple ways in which data can be inaccurate and I think your question sounds like mostly related to data that would be returned to individuals. And I think the data that we would highlight to return individuals with interpretation, if it's just going to be returned, would be those data that are - have the highest action ability and we didn't specify a clear guidance of what that looks like. But that is something that we thought is in the purview of the subcommittee.

And then the other types of data - I think it's a very important question that we have to make sure that those filters that they go through before they actually click, click and say get the data, really the participants are understanding the kind of data that they're getting that what its accuracy might be. We certainly felt very important that everything be done in CLIA fashion, if it's on the front side of that return. And that could be things like (unintelligible). It could certainly be other types of mutations that have high clinical action ability.

(Ian Lipkin): I probably misspoke when I said inaccurate data. What I meant was premature decisions or, you know, or inaccurate assessments of risk based on data. Would it amount to the same thing? I think maybe if there's something about a data safety monitoring port equivalent, you might want to include that in the report.

And the last point had to do with intellectual property.

(Kathy): So, we didn't have extended conversations about intellectual property. And so that's probably an area that we need to have additional conversations about. (Eric), do you remember if we had any - I don't.

(Eric): No. I mean we mostly - we did talk about the possibility of, you know, there were questions about the intellectual property of the researchers doing their research and revealing their -- what they're looking at -- to others. How public were the (unintelligible) and we assumed that the queries were not public (unintelligible). So - but outside of that, I can't remember any more detailed conversations about it.

(Ian Lipkin): Each of these issues has come up in the MoBa, the Norwegian Birth Cohort, which is, you know, now in its 15th year. So I think these are things that you might want to speak with - I can speak with you offline about this if that would be helpful. But we - I could talk with you about the ways in which we handled it.

Francis Collins: Great.

Man 3: That's great. I think in general we would tend to be conservative in being most enthusiastic about returning actionable data that has the most direct impact on health implications and being conservative beyond that. DSMB - obviously

any clinical trials that are initiated within the cohort would have their own data safety and monitoring committees.

And then I think for intellectual property, I can't speak to the cohort you're involved with, but in the United States federal law regulates what -- how intellectual property is allocated, unless otherwise specified. And so in general, it goes to the individual who actually made the discovery that is being filed. And that would be my first pass expectation.

(Chris Wilson): This is (Chris Wilson). I wonder if I think (Ian's) - maybe I'm interpreting his intent by that. And I'll phrase it maybe by saying two things. Who owns the data? Who owns the analytical tools? Is there any corporate nature to that or is it all going to be, if you will, distributed, and then who owns the output of the analyses and how visible are those to others and in what time frame?

Man 3: Right. So that's a - obviously a very layered and complicated question. The data would be in principle stored at the central repository, but would be accessible broadly to the investigator community for free use. So, I'm not sure in the legal sense where ownership would lie, but I think as far as intellectual property goes, or the ability to publish results coming out of the results, it is not envisioned that there would be a central authority that would regulate the use of the data for accepted research purposes.

(Kathy): Yes, you know, and the conversation about ownership of data, you know, we were - there was some intentionality around not talking about who owns the data, because really the data are from the participants. It's the participant's data and we are the stewards of that data and making sure that it's being used in ways that the participants intended and that will be maximally helpful to scientific research. So we tried to talk about where it lives and how it's used, but not about ownership per se of the data.

(Josh): And in terms of outcomes from studies, we did discuss that any data that resulted from use of bio specimen or running some novel algorithm across other classes of data would have to be shared back to the central coordinating center, but with the provisions that there could be embargos on secondary publication of that data to - just like dbGaP has - to enable the investigators to publish, that sort of thing.

(Ian Lipkin): Okay, great. That's really what I was asking. I'm glad to hear that you're saying it's going to fall out through the paradigm that was established through dbGaP. Because there (unintelligible) - it's a given gift that way which I think is actually valuable.

(Cory Barkman): So this is (Cory Barkman). If I could just make a comment and then ask a question.

Francis Collins: Please.

(Cory Barkman): So the comment is I think this is really a terrific program. I think it's a very exciting program and report. I think the committee did an outstanding job. And I particularly - the way the shorthand for the precision medicine initiative had originally been something like, you know, incorporating all of these genomic approaches into medicine, and that was a great idea and very timely. But what turns it into a spectacular idea is working the epidemiology in, working the environmental influences in, considering things like the health disparities. And just reading those sections of the report reminded me of how transformative the Framingham study and the nurses' health study had been, and how this really has the possibility of transforming medicine by adding this whole different, much more complicated element in - the genetics is the easy

part. And I thought that was just a fantastic approach. And I think people don't even know how much they're going to appreciate this.

So my question is - actually sort of curious. What if you're too successful? What if more than a million people want to participate in something like this and, you know, particularly as you try to, you know, be representative and include some groups and balance to their population. I know that you said that at this point, you don't want to go directly through, for example, the patient organizations or foundations. But if there is really a tremendous support from some of these groups that are engaged with patients, might they sort of become partners later in this process? Or have you thought through what it would be if it became two million people?

(Kathy): Yes. You know, I don't think that the working group spent - you know, all of the input we got was really enforcing that you want more than a million -- you want more than a million. And we didn't have, in the working group, the detailed conversation about how to cope with an abundance of success. But certainly, internally, we are - already this morning institute directors met to discuss this report and talking about what would happen if you are, sort of, overachieving.

So we will need to figure out how to do that and also how to make sure that we are getting - we're not filling up all of the available slots with a certain demographic. And so, we want to make sure that we are getting the risk diversity across America and make sure that we are getting underrepresented groups represented in this cohort early.

So, one of the approaches which we didn't spend much time on in this presentation, but we have spent quite a bit of time on in the working group conversations is how to partner with our sister agency, HERSA, who supports

federally qualified health centers who have a majority of folks that they take care of -- patients that they take care of -- that are below the poverty line and have an overrepresented patient of underrepresented groups. So we're really looking to that as a fantastic way to be able to make sure that we have those folks engaged and engaged early on.

But we have clearly a lot of work to do to think about both how do we respond to early success, and also how do we respond if this sort of goes over with a big clunk. And so we're going to have to do some thinking through those scenarios.

(Man 3): Yes, I'll just add to that, (Cory). I think we really struggled with wanted to make sure that in principle anyone can volunteer, because we think that's a great idea for the country. But you could imagine that that could swamp the entire enterprise if truly successful, and you wouldn't end up with the diversity of age and preconditions that we'll need for the cohort to be successful.

So in the report we articulated that we really do envision a very large chunk of at least the first million being composed of people from HPOs where we actually can passively collect healthcare records, follow patients longitudinally, and be confident that we will be well represented from that group. And then view, in some ways, the volunteers who come in entirely on their own as a bonus.

(Sudha): So, hi, this is (Sudha Akeel). So, first I do want to echo what (Cory) said, that this framework is richer, more layered than one might have, you know, guessed or imagined from the initial descriptions and that is wonderful to see. And, you know, I really think it is going to be transformative and hopefully eventually, of course, the whole country is part of the experiment. But, how to stage that is, of course, a big - an interesting challenge.

The - I have two kind of comments that - and questions. One is that I think you have rightly emphasized, (Rick), the issue of thinking about vulnerability and different types of vulnerability -- environmental, social and so on, to various disorders. And I wanted to also just emphasize, if we could, at the same time think of ways of thinking about social resilience, and how people sometimes survive because there might be really important hints, biological and otherwise, about why people who are under, you know, both health and poverty and social distress, still sometimes remain healthy. So that might be kind of an additional - I don't know if the measures are the same or whether some additional measures are needed to capture that. And - or maybe just the way the questions get asked when the data are collected.

But, that's the advantage of getting healthy people as well as non-healthy people into the cohort, and the longitudinal analysis. And the other component of that is how much information is being collected - simple stuff, but questionnaire maybe type stuff about psychological variables. Because as part of that gene environment process is how people respond to their, you know, environment. And a lot of that is temperament. A lot of the genetic studies have shown - at least for psychiatric disorders - but that's the common variable -- that how people react to their environment. And it may not be a lot of work to collect a little bit of additional information that might prove rather predictive over reactivity to environmental or interactions between genetics and environment.

(Rick): So I think these are really great points, and one of the elements that we were very enthusiastic about in discussing the mobile technologies aspect of the program is the ability to do these kinds of surveys. And, you know, with such a high fraction of the population having smart phones, you can imagine doing

these kinds of surveys and collecting the data relatively easily upfront and at appropriate periods throughout the life of the project.

And certainly you're comments about resilience are, I think are completely apt. You know, we tried to emphasize throughout the report both health and disease. And we highlighted the protective mutations in PCSK9 that lower LDL and prevent heart attack. The same applies to behavioral traits. We would be very enthusiastic about studies to look for what are the factors that contribute to resilience or people who have mutations who ought to have particular disease outcomes who don't, and trying to understand why that might be.

Man: Good, great. Thanks.

(Elba Serrano): Hi, this is (Elba Serrano) and I have a question about privacy. On page 82 of the report, you have a discussion of the Freedom of Information Act and the implications of that for even de-identified data. And you also mentioned the requirement for - potential requirement for legislation. Could you elaborate a little bit more on that and whether you see the lack of legislation being a serious impediment for some of the proposed efforts.

(Kathy): (Elba), thank you for the question. I don't have the report in front of me, but I do remember the section. So, the Freedom of Information Act allows us to not disclose private, identifiable information about individuals that we may hold. However, and we have interpreted that to include, for example, genetic information such as we hold in dbGaP.

We - so we feel okay about that and that has been successful up until now. But we have been worried if such a case went to court, whether or not we would be able to defend the position that we have not disclose genomic and other

information. We have had requests. In fact, we've had FOIA requests for dbGaP data, including the HeLa data that we spoke about earlier. We managed to defend against that. (Jason Drey) who's on the call also had a FOIA request for HeLa genome data clearly from people who are not well intentioned.

And so, we would like to have this legislative fix, but I think we will be okay without it. It just gives us an extra sense of protection and security.

(Elba Serrano): Thank you.

(Harlan): This is (Harlan Crolts (SP)) and I just wanted to also just echo the admiration for which everyone accomplished in a very short period of time with an extraordinary amount of actual pressure from various different sources to deliver this complex and thoughtful document. But I'm just thinking about this part about the enrollment and whether - what you guys considered in terms of these innovative ideas to actually make his possible for the - by the time the president leaves office - you've actually got a million people in. And, you know, getting this sense of urgency behind this is kind of moon shot mentality, and can it be done at a high quality level at a rapid pace, using means by which you're really engaging the community. I know that's some big - I've got - I have so much admiration for this, that your seams about engagement are so clear in this document. But I just was wondering what your thoughts were about, you know, what's possible in terms of breaking the sound barrier on really getting up to speed very rapidly?

(Kathy): I think I heard the rapid intake of breath around the table and on the phone with a million people within about - what, 14 months' time. So that would require that we do things in a super-fast - as you say, break the sound barrier. I'm not sure that we'll get to a million people in that period of time, but I think

we do have in mind that we want to make serious inroads into recruitment rapidly so that we don't have a long period of time go by without being able to take this program live.

So I share your enthusiasm. I'm not quite sure that we're - we'll be able to hit that ambitious target. But we'll fire up the engines.

(Harlan): And, you know, it's a stretch. But I'm just kind of thinking conceptually how do you hack this essentially so that, you know, you're basically thinking okay, here's how we would do this. What's available to us to do it in entirely different ways that would enlist the entire country to want to be part of this remarkable effort to, you know, make a big difference in our fight against disease?

(Bray): So, (Harlan), this is (Bray Patrick-Lake). I went dark right as I was about to give some (unintelligible) before, so I really didn't get to speak to, you know, the importance of grass roots, engagement - community engagement and the sense of volunteerism. And so, really while we're talking about recruitment, we're really speaking about engagement as the foundation of this entire study and the success really hinges on that principle.

So, we won't skimp on that part. I feel like the working group has been very thoughtful and if you take a look at section 4, there should be more detail in there for you.

Man 3: So, in the full document on page 35, we reviewed what has actually been done by several sites that have been doing comparable kinds of things. And, you know, the million veteran project has recruited more than 400,000 individuals over a four year period. And so we speculated that with reasonably -- we think readily achievable goals, if you had ten sites doing 25,000 participants per

year, you'd get to a million in four years without any unconnected volunteers coming in entirely on their own.

And, I think the pace that several projects have been moving at suggests that you can - that that's conservative, that you can actually do significantly better than that. But it's a big list of - considering where we're starting today.

(Harlan): And I didn't mean to detract at all from all this, I'm just thinking out loud with you and, you know, what pace it can be done. But it's exciting to contemplate.

(Mike Wells): This is (Mike Wells). It's a significant opportunity to think about how this might be able to impact a variety of abnormalities to the brain. So for example people with a variety of disabilities neuronal diseases, people with neurodegenerative diseases, people with mental illnesses. What thoughts are there about how we might be able to get consent from those people?

Man 2: So, the discussion of consent on people who are decisionally impaired is an area that we did spend time discussing. Clearly the opportunity to collect those individuals in a perspective manner is something that we envision as they were to become decisionally impaired through a variety of circumstances, which could of course include neurologic illness. The actual ways we could recruit those who are decisionally impaired, we - is perhaps a discussion that needs more evolution as the cohort (unintelligible).

Francis Collins: Other questions?

(Ian): So Francis, this is (Ian). I just wanted to get a little bit into the weeds.

Francis Collins: Oh go, sure.

(Ian): And so these points don't get lost in case I don't follow through. The first has to do with questionnaires. Our experience has been that the more complicated the questionnaire, the less likely people will be to complete them and the more of them, the less likely. So it's very important that the groups that have specific questions they want to address, get together and trim those accordingly.

The second point is we have many languages spoken and as we develop questions that we're going to pose, it's going to be very important to do back translations to make certain that you're all asking the same questions of different people. Otherwise, it's going to be very difficult, in retrospect, to know what you know.

And then the - with respect to Biobank, although this - I realize you've decided to start with blood and that makes sense, I think it's very important to talk about things like standardized collection procedures. So in our experience, the time of day when you collect plasma or serum has a large impact on proteomic profile, and certainly an (unintelligible) profile. So this is something you're going to want to consider. Things as simple as whether or not people use EDTA versus citrate tube has a huge impact on proteomic.

So your blood - you're Biobank group is going to have to think about all these very simple things that nobody really thinks about in advance because they think blood is blood, whether it's short minus 20 or minus 80 has a big impact on what you can do with it. So, again, if you want to look at what we did wrong and what we've learned not to do, we'll be happy to share those data with you.

Man 4: Those are all incredibly important points and we really appreciate your thoughtful comments. And the Biobank group is certainly thinking very

carefully about exactly what to collect and how, and we completely agree that things need to be very standardized. And that, of course, is another of the concerns about the volunteers coming in from all across the country, from, you know, and of one, healthcare providers and the ability to really standardize well is a concern for exactly the reasons you articulate. And I have greater confidence that we'll be able to get standardized procedures across HPOs.

Francis Collins: (Ian) thanks for those comments. We are coming within ten minutes of the stated time for this call. I want to be sensitive to everybody's time table. If there were any quick additional questions, we could entertain them, but pretty soon we need to ask the ACD to decide about the acceptance of this report.

(Ian): I'd like to make a motion to accept the report. (Ian).

(Cato): I'll second that.

Francis Collins: That was (Cato)?

(Cato): Yes.

Francis Collins: Great, thank you. Our points of discussion? I am hearing none. I guess we should also then take a vote, and I believe that (Elba) has joined, because we heard from her, and (Chris) has joined. So anybody else who's joined since I last took the role? Okay, then I think I know who is on. And I think we'll do this in the same way we did with the HeLa working group report. I will simply ask for a voice vote and then we'll see if anybody pops up in the other categories besides approval.

So there is a motion on the table to accept the report. All in favor please say I.

Group: Aye.

Francis Collins: Any opposed? Any abstention? All right, well congratulations to this working group for a really wonderful job that you all have done putting forward - I have to say, you know, it's not only a wonderful piece of work in terms of all of the deep deliberations you went through, but also it's a really good read when you have to sit down and read a report that's a hundred pages long, sometimes you really swallow hard and think oh man, I hope this isn't going to be painful. And this one just carried you right along by the very eloquent way that the case was made for all of the recommendations that are in this report. And, again, much credit to all of you and to (Glenn) for making it not just a great report, but one that's fun to read. And I hope it will be widely read by many people who haven't yet seen it, because this is -- as I said at the beginning -- a real inflection point for this study and this will set the tone for where we now are deciding to go.

So, ACD thank you for your vote in favor. That is, of course, means that you are now passing it to me, the NIH director, with your stamp of approval. And when such reports come to me by the ACD, I have the choice of deciding whether to think about it a while or whether to act immediately. And in this instance, I am ready to act immediately.

So, with the understanding that many recommendations here are going to require some flexible interpretation, and I think the working group wrote the report accordingly to allow that because there's a lot of things that still need further investigation. I am happy to officially accept the working group report as now endorsed by the ACD, which means that we are now ready to move from the design plan to implementation of that plan. And I'm delighted to say

on September 17th, that we've reached that point because that was a heck of a lot of work.

But a lot of work lies ahead now and we must, indeed, take your recommendations and figure out how to turn those into reality. I am ready also to make one other announcement, and that is one of your recommendations was that we need a director for the PMI cohort plan. Obviously, we will now, as of the next day or so, put together a plan to carry out a national recruitment for the very best person to come and lead this historic enterprise. But that will take some time to identify the right person for that long-term role.

Meanwhile, we need very much to have somebody to step in and take this on in an acting, sort of interim stage. And I am happy to say that I have approached someone who has now accepted the responsibility of taking this on. Someone who has the right credentials to be able to lead this next part of the effort. And that is (Dr. Josephine Briggs) who happens to be also sitting here in the room. Those of you who know (Josey) will I think immediately recognize just what a great opportunity this is for all of us to tap into her expertise, given the things that she's done over the years. She's a nephrologist, but she's been at NIH in various capacities, and she has had extensive efforts overseeing things like the health systems research collaboratory, which is in many ways got a similar kind of set of goals. You might even think of it as a bit of a pilot for what we're trying to do with PMI. She's also played a very important role as an advisor to P. Cornett and for quite some time she oversaw the CTSA's, the Clinical Translational Science Awards.

So she - and I have the good fortune to work with Josey in many of these areas over the course of the last few years, and I know she has a lot of skills to bring to this, among which are that she is a very good synthesizer -- somebody who really can take information and put it together. She is, at current time, the

institute director of the National Center for Complimentary and Integrative Health, and she will, of course, need to continue to play that role. But she feels that she can, in fact, put a very large amount of her time into the PMI. She's sort of smiling at me, I think.

(Josey): (Unintelligible) says we have to break the sound barrier. (Unintelligible) job description.

Francis Collins: But, anyway. I would just ask all of you please give (Josey) your full support as she steps in here. We will all try to do everything we can to make that successful.

(Josey): I may call some of the members of this body for amplification of your very thoughtful comments.

Francis Collins: Yes, indeed. And again, working group, you have been so helpful. We are, I guess, not asking you to continue your workshops every three weeks, but we are asking you to continue to be available for the kind of consultations that (Josey) just mentioned.

So, I think with that, and with now much to do in front of us to take these recommendations and move them forward, under the understanding that I would suspect the President of the United States will be very unhappy if he walks out of the White House and this cohort is not well along, including having enrolled a fair number of people. A million, well that may be a bit of a stretch. But, certainly some percentage of that that's in the double digits would be awfully nice to be able to point to. And I think we can do that. And (Harlan), I appreciated your comments about let's - was it (Harlan).

(Kathy): Yes, I've got it (unintelligible) the box.

Francis Collins: Break the mold, think out of the box. And we have been thinking out of the box about how to do this in a fashion that is not just the same old FOA method, although we're going to be using some of those quite a bit as well. So, yes, watch this space. You're likely to see some things that maybe NIH isn't usually been using as far as a means to get this off the ground. We're ready to try some new experiments here.

So, I think that means we have -- at 2:56 eastern time, pretty much reached the point that we hope to have reached. Let me remind the members of the ACD to complete your conflict of interest forms - always something to pay attention to. They're available on the share file site and email those back to (Gretchen Wood). And let me also remind you to check your calendars, ACD members, and be sure that you have down December 10 and 11 for our next face-to-face meeting here in Bethesda. We will have a very full, busy and interesting agenda, as we always do. And there will be probably quite a bit to say by then about where we have gone with precision medicine, now that you have given us this charge. So...

Man 5: Francis?

Francis Collins: Yes?

Man 5: Before we close, any thoughts on NIMH and (Tom Insel)?

Francis Collins: Well, there's a lot of grieving, that would be one kind of thought. But we understand the pull that (Tom) has felt to go into a new direction and for him to join up with Google Life Sciences is going to be both a lot of fun for him and great for them. So he goes with a little bit of grieving, but also a lot of celebrating for what he's going to accomplish in this next chapter in his very

distinguished career. And we will be setting up a national search very quickly to look for a replacement. Meanwhile, the acting will be (Dr. Cuthbert), who I think will handle this quite nicely and for a period of months until we can bring the new director to that role.

You know, we have a lot of talented people who run institutes. Most of them ultimately decide that this is not a life sentence, and they do something else. And (Tom) has made that decision and gosh, he's been such a phenomenal leader for all of us, and we expect we can still find him when we need advice. But we will go out and find an equally wonderful leader, if that's possible.

All right. I think that does it. Thank you, everybody, for spending two hours with us in this rather important and I think rather historic discussion. And we'll have much more to say about where it goes next. With that, we are adjourned.

(Group): Thank you.

END