National Committee on Vital and Health Statistics
Meeting of the Full Committee

Transcript
April 1, 2021   10:00 a.m. – 5:00 p.m. ET

Virtual

SPEAKERS

NCVHS Members		
Name	**Organization**	**Role**
Nicholas L. Coussoule	**Horizon BlueCross BlueShield of New Jersey**	**Chair**
Sharon Arnold	**DHHS**	**Executive Staff Director**
Rebecca Hines	**NCHS**	**Executive Secretary**
Debra Strickland	Conduent	Member
Denise Chrysler	University of Michigan School of Public Health	Member
Denise E. Love	Individual	Member
Jacki Monson	Sutter Health	Member
Jamie Ferguson	Kaiser Permanente	Member
James J. Cimino	University of Alabama at Birmingham	Member
Tammy Banks	Individual	Member
Margaret A. Skurka	Indiana University Northwest and Principal, MAS, Inc	Member
Melissa M. Goldstein	The George Washington University	Member
Richard W. Landen	Individual	Member
Valerie Watzlaf	University of Pittsburgh	Member
Vickie M. Mays	UCLA	Member
Wu Xu	Individual	Member
NCVHS Staff		
Name	**Organization**	**Role**
Maya Bernstein	ASPE/OSDP	Staff
Lorraine Doo	CMS	Staff
Rachel Seeger	HHS Office for Civil Rights	Staff
Marietta Squire	NCHS	Staff
Geneva Cashaw	NCHS	Staff

Presenters		
Name	**Organization**	**Role**
Daniel B. Jernigan	CDC	Deputy Director for Public Health Science and Surveillance
Elizabeth Pathak	COVKID Project	Director
Seth Spielman	University of Colorado Boulder	Chief Data Strategy and Analytics Officer
David Van Riper	Minnesota's Population Center	Director of Spatial Analysis
Betty Bekemeier	Washington School of Nursing	Endowed Professor
Alex Stokes	BU School of Public Health	Assistant Professor in Global Health
Glen Mays	University of Colorado	Professor and Chair of Health Systems Management and Policy
Michael Soto	Georgetown University	Professor of Health Systems Administration and Population Health
Ninez Ponce	UCLA Center for Health Policy Research	Director
Mark Cullen	Stanford University	Professor of Medicine

Call to Order/Roll Call

Rebecca Hines: Good morning everyone. Welcome back to day two to our members and to the public and our committee staff. It is good to see you all again. I hope you had a restful night, ready for day two. This is day 2 of the Spring meeting of the National Committee on Vital and Health Statistics. My names is Rebecca Hines and I serve as executive secretary and designated federal officer for the Committee.

Let's move right in to roll call. We will go in the same order as we did yesterday, beginning with our chair.

Nick Coussoule: Good morning all. Welcome back. This is Nick Coussoule. I am a Senior Vice President, Enterprise Business & Technology Solutions at Horizon Blue Cross Blue Shield of New Jersey. Chair of the Full Committee and I have no conflicts.

Rebecca Hines: Deb.

Debra Strickland: Hi, I am Deb Strickland. I am a member of the Full Committee and a member of the Standard Subcommittee and I have no conflicts.

Rebecca Hines: Denise Chrysler.

Denise Chrysler: Good morning. My name is Denise Chrysler. I work for the University of Michigan School of Public Health and with the Network for Public Health Law. I am a member of the Full Committee, I serve on the Privacy, Confidentiality and Security Subcommittee, and I have no conflicts.

Rebecca Hines: Denise Love.

Denise Love: Denise Love, retired public health data consultant. I am a member of the Full Committee and a co-chair of the Standard Subcommittee, and no conflicts.

Rebecca Hines: Jacki.

Jacki Monson: Good morning. Jacki Monson, Vice President, Sutter Health. I am a member of the Full Committee, co-chair of the Privacy, Security Confidentiality Committee, and no conflicts.

Rebecca Hines: Jamie.

Jamie Ferguson: Good morning. I am Jamie Ferguson. I work for Kaiser Permanente. I am a member of the Full Committee and the Subcommittee on Standards, and I have no conflicts.

Rebecca Hines: Jim.

Jim Cimino: Jim Cimino. I am Professor of Medicine, Director of the Informatics Institute, University of Alabama at Birmingham. I am a member of the Full Committee, member of the Standard Subcommittee, and I have no conflicts.

Rebecca Hines: Margaret.

Margaret Skurka: Hi, I am Margaret Skurka. I am a member of the Full Committee and I am a member of the Standards Subcommittee. I am a Professor Emeritus at Indiana University. Currently own a consulting company in suburban Chicago, and I have no conflicts.

Rebecca Hines: Thank you. Melissa.

Melissa Goldstein: Good morning. I am Melissa Goldstein. I am a professor at George Washington University. I am a member of the Full Committee and a co-chair of the Subcommittee on Privacy, Confidentiality and Security and I have no conflicts.

Rebecca Hines: Rich.

Rich Landen: Good morning. Rich Landen. Member Full Committee, co-chair of the Subcommittee on Standards. I have no conflicts.

Rebecca Hines: Tammy.

[inaudible] Greg, we have the issue again with Tammy's audio. We will circle back with you Tammy.

Valerie.

Valerie Watzlaf: Good morning. I am Valerie Watzlaf. I work for the University of Pittsburgh as Associate Professor and Vice chair of Education in the Department of Health Information Management. I am a member of the Full Committee and I serve on the Subcommittee of Privacy, Confidentiality and Security. I have no conflicts.

Rebecca Hines: Vickie.

Vickie Mays: Good morning everyone. Vickie Mays. I am a Distinguished Professor at UCLA in the Department of Health Policy and Management in Psychology. I am a member of the Full Committee. I am a member of the Subcommittee on Privacy, Confidentiality and Security and I have no conflicts.

Rebecca Hines: Wu.

Wu Xu: My name is Wu Xu. I am a retired state public health informatics director. I am a member of the Full Committee and no conflicts.

Rebecca Hines: Thank you. That is the membership. We will circle back with Tammy Banks when she is able to reconnect. For staff, Lorraine.

Lorraine Doo: Good morning. This is Lorraine Doo with the Health Informatics and Interoperability Group in the Office of Burden Reduction and Informatics at the Centers for Medicare and Medicaid Services and lead staff to the Standards Subcommittee.

Rebecca Hines: Rachel.

Rachel Seeger: Good morning. I am Rachel Seeger. I am senior advisor with HHS Office for Civil Rights. I serve as lead staff to the Subcommittee on Privacy, Confidentiality and Security.

Rebecca Hines: Great. Any other staff want to say good morning? Alright, Tammy, looks like you are back on. We will have to circle back to read Tammy in in a little while. Nick, help me remember that. Looks like she has some audio challenges this morning.

Just a reminder before we get started, that public comment is scheduled for 4:45 today. So we are going to try and make that happen on time. The afternoon after the Expert Panel is a little fluid, but we will aim for that. Thank you.

Welcome Remarks/Agenda Review

Nick Coussoule: Can we pull the agenda up for today? Okay. We've got a busy day today, an exciting day. We will start off 10:15 with our Subcommittee on Privacy, Confidentiality and Security, talking about their proposed project and other updates that they would like to give to the full committee.

Then we have Daniel Jernigan, who is deputy director of public health science and surveillance at CDC, talking about the CDC data modernization initiative.

We will take a break for lunch, and then we have a fantastic panel, or a couple panels, this afternoon in regards to COVID-19, capacity gaps, quality data collection of race and ethnicity data. That will be moderated by Vickie Mays and Denise Love, with lots of different invited experts. You can see more of them on the rest of the agenda.

Then we'll take a break, and then we have a couple of wrap-ups. One, we're going to spend a little more time on our report to Congress on going through some of the detailed texts in the first few sections. We will spend a bit of time going over our workplan, where our subcommittees will be updating us a little bit on some of the scheduling, or potential scheduling aspects.

We'll finish with public comments and close, hopefully, on time at 5 o'clock. With that said, I will turn it over to Melissa and Jacki. I'm not sure who is going to start us off on our privacy, confidentiality and security subcommittee.

Subcommittee on Privacy, Confidentiality and Security

Jacki Monson: Good morning. I'm going to lead us off, and Melissa is going to chime in as needed. So let's go ahead and get started. My plan is to cover our slides first, and then we can discuss the scoping document. Today I plan to give you the background, the project goals, the phase plan, proposed timeline, and discussion, specifically related to our desire to focus on information security.

We thought it would be helpful to just talk about NCVHS's charge, which is studying and identifying privacy and security and access measures to protect individually identifiable health information in an environment of electronic networking and multiple uses of data.

Here's some of the background on why we want to focus on information security. There's lots of challenges in safeguarding individually identifiable information that has become, I think, increasingly complex as we move into the electronic environment over the past few years. We've had a lot of opportunities in the areas, rise in hacking, and information security related incidents, as well as insider threats, which I actually think are more complicated with the remote environment that we all moved into very quickly last year and are still in today.

The other areas, advancing interoperability, we have new technology, we have use of big data, artificial intelligence, are all new opportunities which are good for the industry, but also create opportunity with meaningful engagement for our consumers, families, and communities in advancing health and healthcare, but also create information security risk.

Privacy and security are linked to all of the ways in which the information about the individuals is collected, analyzed, and used in our obviously increasing digital environment. Then we've had a very significant increase in the growing number of ransomware attacks and targeted nation-state attacks illustrates the vulnerability of this information. You'll see a couple of news articles on one of the next slides, which there's been a very strong focus since COVID started on the desire to get COVID research data, really any data around nation-states, which is part of the reason why we wanted to discuss this with you today.

More background, obviously consumer trust in the privacy and security space is extremely important. Balancing that with the emerging technologies can help individuals make more informed decisions about their care and how they want technology used, specifically to that.

When breaches of healthcare information occur, they can be serious consequences for the organization and for patients alike, including reputational and financial harm, and loss of trust for that patient. Poor security practices heighten the vulnerability of patients' information in health information systems, and the risk just continues to increase.

Here are some of the headlines that just highlight some of the areas. I think every single day there's new headlines surrounding this, so these are probably out of date from just specifically two today, because there's literally ransomware, malware, cybersecurity issues, on a level frequency just continues to increase.

This data is from the Office for Civil Rights, and this just illustrates from 2016 through 2020 the breaches impacting over 500 individuals or more, you'll see that these are specific to hacking and IT incidents, and they just continue to go up. You'll see in 2019 and 2020 they are very similar levels. Certainly, we'll look for 2021 data which we expect to continue to rise.

This is another example of HHS OCR's data on breaches affecting more than 500, and this one is specifically with respect to email accounts, which would fall under information security incidents.

The goal of this project and what we're recommending is a focus on privacy and security stewardship and framework guidance, which we've been doing for the last two decades. We've made a number of reports suggesting enhancements to protections of use of health data, stewardship framework, and have really focused in the last couple of years on what things have fallen beyond the HIPAA guidance and the growing ranges of uses of really data.

In response to the growing number of cyber threats impacting the healthcare industry and the risk to privacy and security of the individuals' information, we would like to take on a project to examine solutions for improving the security posture of the healthcare industry both at the federal, state, local, and tribal organization level.

One of the project goals is to identify and describe the changing security landscape and risk to the privacy and security of individually identifiable information held by the healthcare industry, policies, practices, and technologies. The goal would be to lay out integrative models for how best to secure

individually identifiable information, while enabling useful uses, services, and technologies; formulate recommendations for the Secretary on the actions that HHS might take; and prepare a report for the Secretary.

The phased approach. The first phase we're suggesting is to conduct an environmental scan to explore key security challenges and opportunities for securing individually identifiable information. This will likely include the current security landscape, evolving technologies for safeguarding this information, existing frameworks -- the HIPAA security rule, the NIST standard -- talk about artificial intelligence, big data, and other expanded uses and users; include the internet of things, medical devices, applications, and emerging technologies, and then of course, the novel approaches to enhancing enterprise-wide security.

The approach, we plan to accomplish the environmental scan with a virtual hearing and background research to learn from a range of federal agencies, academics, technologists, and thought leaders in this area. One hearing will be conducted along with background research. Participants may include representatives from federal sources. You can see the list there, there's quite a few of them that are involved in information security. And then of course private sector stakeholders, including academics, CIOs, technologists, and thought leaders. The deliverable would be a report of the environmental scan findings.

The second phase, based on what we learn from the environmental scan, will be to develop models and illustrative future scenarios, laying out assumptions, and identifying areas of uncertainty. The approach would be to develop models and identify potential policy, practice, and technology solutions, and then the deliverable would be a draft security framework, including benefits, levers, and relationships to policy alternatives.

Phase three would be to prepare recommendations for the Secretary of HHS. It might include a framework, best practices, levers that HHS can apply, and legislative mechanisms such as enforcement. The approach would be of course the preparation of the letter, and then the letter to the Secretary be drafted and approved by this committee.

Phase four would be to prepare the report for the healthcare industry and data stewards in the use of healthcare data reflecting a security framework and policy and practice recommendations. This would be modeled on earlier NCVHS stewardship primers and frameworks that we've previously done.

Our proposed timeline, right now, we're currently proposed to be in phase one for the rest of this year, then the plan would be towards the need of this year to be in phase two, which is the development of the draft framework, and then phase three would be prepare and approve the letter to HHS, which would be 2022, first, second, and possibly third quarter. And then obviously phase four will be turning into learning and recommendations, which we would think could happen at the same time as phase three, in the beginning of 2022, likely going through Q3 of 2022.

That's really the premise of it, so before we get into the scoping document, it would be great to hear from committee members on your thoughts and feedback for us, and also to my co-chair Melissa, anything that you'd like to add.

Melissa Goldstein: Thanks, Jacki, that was a really comprehensive and fabulous presentation. I'd like to thank Rachel Seeger particularly for helping us put the material together, and other staff members too, including our two interns, who have been a great help along the way. I also want to say that the list of

both areas that we want to cover in the environmental scan and the people that we might hear from during the hearings is meant to be illustrative. So we listed all of our federal agency alphabet soup, but we also really are hoping to know what's out there and to get ideas from state and local and tribal officials as well, people in the private sector, all over the private sector, so that -- and consumers and patients -- so we want to hear what's out there and learn what's out there, before we move onto the next stages of the project. And we mean that comprehensively.

Jacki Monson: I know it is early, but any thoughts and feedback on the presentation before we move into the scoping document?

Nick Coussoule: This is Nick, let me start. One of my concerns -- that's the wrong choice of words, but one of the challenges is this is such a large scope of work, potential, and there is to think about how we try to to some degree get enough information that's big enough and then figure out how to, to some degree, skinny it down to something we can manage. How are you thinking about that dynamic of such a large and complicated effort, to try to pack in?

Jacki Monson: I think what we are thinking about is doing the hearing first, because this area is so broad in scope, inviting the experts to try to drive what the focus is. We think that there will be commonalities and trends among what they tell us, that will allow us to narrowly focus it. Then we're also working under the premise that not a lot of folks even on our subcommittee have a lot of expertise in information security, so it would be good to hear from those experts first, and hear kind of what's going on in industry, what's going on in government, and how we can help as a committee. I think today it would be hard to narrow the scope, but after the hearing we would have a strong desire to do that, to focus our efforts on what would be most meaningful.

Jamie Ferguson: First of all, I want to say I strongly support this proposal overall. I think this is an important activity for the committee to undertake. But I would also observe that specific technology threats, approaches and solutions, in this space, are changing too fast to be addressed by a project of this kind. So I would urge you to focus this more on policy and governance and seeking to ensure appropriate approaches to things like threat information sharing, rather than focusing on specific technology, threats, approaches, or solutions.

Jacki Monson: Thank you for that feedback.

Rich Landen: I guess my question is very similar to Nick's and Jamie's. In the overview you talked about state-sponsored hacking, and I don't know that the assumption that we'll be experts at that, that NCVHS will be experts at that. This could be taken overly seriously, but I guess, like Jamie, I'm kind of wondering where our sweet spot is. What's our role in this? Obviously, we're not going to identify specific solutions, so again, the project is very attractive, I think there's lot of value to it, I'm very supportive. But I'm not sure where we should be focusing or stated the other way around, what pieces of this project would be best left to other technical experts.

Jacki Monson: Yes, it's a fair comment, and that I think goes into narrowing the scope once we've done the hearing, to figure out what will be the most meaningful contributions that we can make. I think today, without doing the actual environmental scan, without doing that hearing, I'm not sure that we would want to make that decision automatically on what would be the most meaningful. I think we want to let that information drive it, and then identify and obviously come back to the full committee in discussion about where we think we can make meaningful contributions to this very concerning area where I think we do have the ability as we talked about in the scope in the very beginning of the

presentation, we do have the opportunity as this committee, and part of our scope, to I think make a meaningful contribution. It's just identifying what specifically that will be, and we'll let the hearing drive that.

Rich Landen: Thanks, that reasoning addresses my question head-on. Much appreciated.

Denise Chrysler: I fully agree. I am a member of the subcommittee. I fully agree with Jacki that we need to have a hearing to be able to hear the breadth of the topic and narrow accordingly, but I also resonate with what Jamie talked about and how this work might enhance other work that we are already involved with. For example, in just thinking about this afternoon's expert panel, and deficits in data collection related to race, and deficits in reporting and willingness to share information with the government or others, and the fear of data being accessed inappropriately within an organization and the governance and the policy and procedures and the ways data are protected from misuse and abuse, and inappropriate access that come from within an organization, and not just a hacking type of situation.

Jacki Monson: Thanks, Denise, for that comment.

Melissa Goldstein: I want to add back up entirely what Denise just said, and also note that one of the things that Micky Tripathi mentioned yesterday in his presentation, and that we discussed a certain amount, are the idea of vaccine passports or certificates that can be used to, I guess, prove that people have been vaccinated in order to get a benefit or access, or something. The more information that we are asking that people share, the larger the trust issue might become, and I think it underscores the incredible importance that we begin to do this work quickly and try to see where the needs are. Because we know that we are not -- that our current policy levers are not actually working to the extent that we want them to.

We can see that from the data on breaches, and we can see it with the hacks on healthcare information, like we've never seen, to the extent that we've seen it recently. So I do want to thank Denise for her comments and underscore and point out that the federal government, and at least HHS, is looking at these issues as well.

Wu Xu: First, I want to say the scope of work cannot reasonably exclude public health, what I understand is good for the security protection part. But for multiple use for sharing, I hope the subcommittee can come up some advice or recommendations how to guide this multiple use sharing between the public health with researchers, now the health systems become to do the research, the quality improvement. So how do we share deidentified data? It's really a challenge for public health. It's just my thought. I don't want to expand your scope of work too much.

Jacki Monson: Thank you for that. We will think about it.

Nick Coussoule: Just a couple other comments from my side. And then we'll go to Denise. I'm frankly a big supporter of this effort. I also happen to be an operator, so as Jacki, we live this every day. There's an interesting -- there's a natural tension that's becoming more and more apparent every day between what I'll call data liquidity and sharing, so how much more we can make available for all the right reasons, at the same time that creates a lot of different -- much higher risks and different risks, in regards to potential exposures and breaches.

So I'll be really curious as to how we can surface a lot of those types of questions, and then ideally frame something up and come to the things that might be good policy levers that would let us do that.

Security, I'll editorialize a little bit, security can often be viewed as a binary kind of question. Something is either good or bad. But we don't live in a binary world. We live in a risk world, where we have to measure some of those things. I think I'm excited about the opportunity to hear from lots of people in regard to both the detailed challenges and the policy challenges we have today. I do agree that getting into the very technical details, that's not really the purview of us. There's others that are doing that already. It'll be interesting to see how we can get enough of that information to help inform us as part of our deliberations. I'm excited about the effort and opportunity.

Denise Love: Yes, thank you. I totally agree with the comments that preceded me, and it's a little redundant, but my notes here, as I heard the presentation is, I think a broader view, like Wu says, for facilitating our sharing of data between public, private sector, any of them within the public sector. So maybe part of this package is some guidance on risk assessment, because I think it's easy to say no, and that's 100 percent secure, if you're making a decision whether to exchange or share data, but that doesn't serve the public good. So how to balance that out with some reasonable risk assessment guidelines.

Jacki Monson: Thank you for that, and we can definitely discuss that as a subcommittee. I think that will come into more of the recommendations part, depending on what we hear. But risk assessments, we know, is a challenging area, and I think anything we decide to do is going to be hard to create prescriptive guidance on how somebody subjectively evaluates risk. But I've noted it, and we'll talk about it as a subcommittee and see how we can try to achieve what everybody's looking for.

Rich Landen: Jacki, in one of your opening slides, you mentioned that this is building on privacy and security recommendations that the committee has been making for the last couple of decades. In part my question is curiosity, but in part it goes to informing us later on down the road, as to how we craft our recommendations. So the question is, do we have a sense, in general, of how, what kind of uptake we've had over the last two decades? If we recommend something, does it pretty much automatically happen, or do we need to work for it?

Again, as I said, in part it's curious, but in part, it's down the road packaging the recommendation to make sure the uptake is maximized.

Jacki Monson: I don't know the answer to that question, because I haven't been on the committee for two decades. I can tell you that since I've been on the committee, I think our work and our recommendations on beyond HIPAA and deidentification have been very well respected and used. And so I would expect that to be the same for any kind of recommendations that we come up with in respect to this area. Not only is it a hot area, but I think people are grappling with what can we do to address the concerns and the risk environment, and so that's what we'd we be looking to narrow in scope.

Vickie Mays: My question is a little bit like in the same domain as Rich's, and as I listened to the discussion, one of the things I wondered about is the priority of who we're actually going to try and answer the greatest good for, and whether -- because what may be an industry problem may be very different for HHS. So I guess the question really is is there guidance within HHS about what their issues are, so that we can make sure that as we invite people to the hearing, that we're balancing the needs for within HHS and industry. So I don't know if there are -- and this may be a Rebecca question -- if there are partners within that we should be inviting to the table early, for them to give us a sense, so that our invitation list is a good balance.

Jacki Monson: I will comment a little bit and then Rachel, if you are on, I think it would be helpful to hear from you. The Healthcare Cybersecurity Taskforce created a report. Many of you know that I was a member of that group, and that report was delivered in 2017. As a result of that, they established the Healthcare Sector Coordinating Council, which is public and government, and then they also had created CISA, which sits within HHS, which is specifically focused on cyber response.

So I think those are two areas where we would obviously want to invite them to the table, to see what needs or progress has been done on that cyber taskforce report from 2017, as well as what CISA has specifically been focused on, and learning about the barriers to their success that we could meaningfully help with. I'm sure there's other agencies, so I'll be quiet and let Rachel talk and add anything that she wants to to respond to your question, Vickie.

Rachel Seeger: What Jacki was saying, there are a number of federal partners, while OCR focuses more on the HIPAA space, and ASPR focuses more on health ecosystem overall. We also want to invite our partners from CISA and other places around government, Homeland Security, to make sure that you all are hearing from a wide view of experts who are addressing the security threats to healthcare overall, not just in the fast lane.

I also just want to add that we're going to try to keep it technology neutral. I hope that addresses some of Jamie's concerns and concerns that others mentioned. One of the issues that we're seeing with this rise in hacking, IT threats, that are affecting the healthcare industry, is because there are so many backdoors open. In the HIPAA sphere, entities and their business associates have done a great job in protecting and locking down the EHR, in order to qualify for the incentive vendor (audio interference), but really what we're seeing is an overall issue with enterprise-wide security that's affecting the entire cybersecurity posture of the healthcare industry, where we have bad actors as well as inside threats, that know that these backdoors are open and vulnerable.

So hopefully we can hear from a wide variety of folks, people who we hear from all the time, but also consumer perspective and other federal partners, other thought leaders. We've done a lot of work on deidentification in the past, have heard from a lot of really good people with thoughts about lifecycle data management and lifecycle data security. So hopefully we can have a thorough environmental scan and a really good hearing that can help inform this subcommittee and the full committee on how to narrow this topic down and get it in the lane for NCVHS where it's going to be meaningful to the department and to the industry.

Vickie Mays: The only thing I'd request is to make sure, because this is where I think the new sweet spot would be, is to start also looking at public health. I think we've been very embedded in cybersecurity issues for healthcare, but it's starting to be an issue in public health, and it's starting to be an issue of public trust. To the extent that we can uncover those kinds of cybersecurity concerns, I think it'll put us in a nice balance. Thank you.

Tammy Banks: Thank you. Really supportive of this request. And like Jamie said, the players in this space with data exchange is increasing rapidly. What are your thoughts in regard to including looking into the patient consumer perspective of this? I know they're not a covered entity. However, more data is being generated by the patient, either through wearables, through medical at home, through telehealth, through requesting price. Is that going to be part of the scope, expanded? I know we're talking about narrowing, but again, it's a growing area that really needs some protection.

Jacki Monson: That has to be part of the scope, right? Because part of, as we talked about in the presentation, patient trust and understanding of how their wearable, where that data goes and how it's used, so not getting too much into the data piece of it but the security of it, I think that's within scope, and we would expect to hear from and try to recruit both the consumer patient advocate perspective, because I think we have to have that and be able to evaluate what would be most meaningful. So that's certainly our desire.

Melissa Goldstein: I would note that Sharon Arnold has been able to join us now, and give Sharon an opportunity to comment, particularly on what's going on in the federal government at this point with security issues. If you'd like, Sharon.

Sharon Arnold: Thanks very much. I joined very late into this discussion, so if somebody could just kind of summarize for me, that would be really helpful. I don't think I got the full gist of what's being discussed. Apologies for joining late.

Melissa Goldstein: We have just been talking about the focus that we're planning for our security project and what is out there right now, how we'll plan to scale down the broad scope that we're focusing on now, what else is going on in the federal government, our federal partners that are also working on this. Rachel and Jacki have given us a good sense of the HHS work. I believe there are interagency groups on this, as well, which Jacki may have been participating in. So really just an opportunity to comment, if you'd like.

Sharon Arnold: I think that this is an area where the committee has excellent expertise and can really complement the work that the federal government is doing or is planning, and it sounds like Rachel is well aware and incorporated into that work, and I think it's terrific.

Melissa Goldstein: Agreed.

Nick Coussoule: Tammy, did you have something additional? Any other questions, comments, from the committee? All right, I'll turn it back to Jacki, then.

Jacki Monson: I would like to pull up the scoping document. I think we're looking for, on the scoping document, not necessarily typos, things like that. We're looking for big-picture comments, flaws, opportunities within it. So let's start with the background, and please just raise your hand as we go to chime in on anything that is said, and then obviously, based on comments we've already received, we'll look at the scoping document offline and look to incorporate the suggestions of the committee.

First, let's start with the background. Basically, we're just scoping up what I talked about specific to the presentation really around privacy and security and access measures to protect individually identifiable health information, in the environment of electronic networking and multiple uses of data. Then we go into the specific challenges around this area that we've talked about in the presentation and our discussion today, and tried to not narrow the scope of this. It's pretty broad, as you can see, of all the areas where we think there might be specific opportunities.

We can scroll down just a little bit. So that really covers our challenges. Then we go into consumer trust. Once we're past consumers' trust, we discuss specifically the frameworks over the last two decades, so this, Rich, will go to some of your thoughts or questions, that I maybe didn't answer. Specific areas in 2008 of things that we focused on, as well as in 2012. Then we sort of wrapped that all up in the background, to just talk about the cyber threats and impact to the healthcare industry.

So I'm going to pause there and see if there's any feedback specific to the background that we haven't already discussed that people would like to note for us to consider enhancing the background to include.

Nick Coussoule: Jacki, this is Nick. The only that I might -- and I think it's sort of referenced in the next-to-last paragraph, where it indicates that recently NCVHS addressed issues that fall beyond the scope of HIPAA, and I think there's a couple good pieces of work that were done by the committee, both beyond HIPAA and some of the deidentification work which I think could be an interesting input into helping us also identify the right folks to bring in. There's some really good work done by the committee over the last five or six years or so.

Jacki Monson: That is a great suggestion, Nick.

Denise Love: I am really liking the scope of work. And I just have a question, and I hate to bring up the COVID word, but we had that workshop that had some good materials and referenced some of the challenges around COVID, so I'm assuming that would be some of that body of work would be rolled into this scope, as well.

Jacki Monson: Denise, are you specifically referencing the hearing that we had in September?

Denise Love: Yes. And they touched on some excellent dilemmas and issues there.

Jacki Monson: We can certainly add that. Our plan with respect to the hearing in September is to actually create a formal list of recommendations to send a letter to the Secretary on, but we can certainly look at that to see if there's anything else that we could enhance this background to include things from that hearing.

Sharon Arnold: This is Sharon. I would support that as well. I think that would be really helpful.

Nick Coussoule: Thank you. Rich?

Rich Landen: There is one thing that hasn't been mentioned yet that I'm almost hesitant to raise, because it's got a huge potential for taking us down a rabbit hole, but I also think that we need to at least talk about it and see whether it fits or not, and that is back in, I believe it was, the 2016 EHR certification rules, the Office of the National Coordinator made it a requirement for EHRs to allow patients to designate third-party app developers for their apps to get information to download the patient record. One of the industry concerns back at that time was that bad actor apps could take advantage of underinformed consumers and do bad things with that data.

I'm wondering if as part of the environmental scan we might take a look at that, and I'm hesitant in some ways, because it's not really a security issue, per se, but it certainly goes to affect trust of the system, so I would be very interested if we could include in some of the environmental scan activities finding out whether that industry concern of whether this patient designation app coming out of -- well, I won't pick a country, but some foreign country -- and essentially then that information is used to steal the patient's identity. So thoughts on that?

Jacki Monson: I don't have any thoughts other than that's a huge opportunity. I think we can't predict what they're going to tell us at the hearing, but certainly we can invite both patients and consumer advocates that will provide that perspective on their reaction to something like that, and I think we can

frame our questions that we ask specific to that hearing, to try to get that information out. But I think beyond that, we'll just have to let the hearing in my opinion kind of drive what we're going to do and what we hear from it. Rich, I don't know if you have other perspective on this.

Rich Landen: Jacki, to be clear, I guess my first question is has that fear materialized? If it hasn't happened, then maybe it's not a major issue.

Jacki Monson: I can tell you from a practitioner's standpoint that it's definitely a dialogue that I hear from patients on a regular basis. So I think it has materialized. To the degree that it has materialized, I think it's going to depend on the audience. For example, you might hear something very different from millennials than you're going to hear from a consumer that is a little bit older, who is more informed and cares about their privacy rights more than necessarily a millennial, who we often hear they just want the functionality within the app and aren't really focused on the other pieces of it. I think we're going to hear that, and I would be surprised if we didn't hear it in the hearing, and we'll sort out how specifically how that information might help us with some recommendations.

Rich Landen: Thanks.

Melissa Goldstein: I would also add that there are multiple surveys that have been conducted that do show that concern. It's actually a much broader concern in general, right now, in terms of the proliferation of apps and vaccine hesitancy and many areas within the healthcare sector, and I'm not just talking about industry. I'm talking healthcare sector-wide, that is receiving attention, but not necessarily solutions yet, and policy recommendations. And I think it's important that we explore it more.

Nick Coussoule: I think the, I'm trying to think of who mentioned it earlier, I don't know if it was Rachel or Jacki or Melissa, but getting the consumer perspective as part of this exercise will be really important, because the whole expectation of privacy question, and understanding even of security and data risks in themselves, we touched on a little bit on that in some of the previous committee work, but I think just given the sheer proliferation of how that's happened over the last handful of years would be really important, to make sure that perspective is included in the work we're going to do.

Lorraine Doo: Nick, this is Lorraine. You may also want to talk to the folks at Medicare with the Blue Button and the consumer app that they have, because I don't know that there's millions of people using it; there's millions of people that have access to it, because we don't know what the data is, but all those apps are vetted pretty carefully, but also the interoperability and patient access final rule goes into effect for Medicare and Medicaid and the qualified health plans in July, and again, we won't have that data. But depending on when this happens would be another potential source for inquiry after that gets implemented later in the year.

Nick Coussoule: I think that is great. Sorry, Jacki, I don't mean to interrupt, but I think it's also, we've got several of us on the committee who are practitioners as well as either policy or theoreticians in some of these things, so that's really important in our day-to-day work.

Jacki Monson: Thank you, Lorraine, for that. I jotted a note down, so that's one area that we'll discuss.

If we can move into the goals area of the project, if we can scroll down a bit on the screen. We've talked a lot about the goals. It really is to just look at the range of security challenges affecting all of healthcare, business partners, consumers, et cetera, so that's really what our goal is. And as I

mentioned before, when you see in the points down here is we've left it intentionally very broad because we really want to let the hearing drive what we focus on, and we believe that we will hear trends and areas where we can focus to most meaningfully contribute to this area. So that's really what our plan is for now.

Again, we plan to identify and describe the challenges, lay out the models that could potentially be used. As I think Rachel clarified, our plan is to not specifically focus on technology, to be agnostic in that area, and focus more generally and broadly, and then obviously, formulate recommendations to the Secretary, and then prepare a report.

So I'll pause there and see if there's any concerns or areas for opportunities to enhance the specific goals that we have laid out in this document.

Rich Landen: Nick, a question. At point 4 here, we say prepare a report for the Secretary. If I recall correctly, the slide said that report would be for industry. In my mind, I suspect, they're pretty synonymous, but is there anything intentional about Secretary versus industry for the report?

Maya Bernstein: Yes. This is Maya. We advise the Secretary. We don't advise industry. Although our reports are available to the public, and industry certainly looks at them and absorbs them, our report is always to the Secretary. Our recommendations can only be made to the Secretary. So that's why.

Nick Coussoule: I think maybe, Rich, you were referring to, in the sequencing event of the workplan components that showed that kind of two paths at the bottom? I'm trying to remember back to the document itself.

Rich Landen: I think it was phase four, was talking about a report to industry.

Rebecca Hines: I wanted to weigh in on the same, because I had a similar question. Sometimes what the committee has done since I've been on board is we'll get the report done, and then you'll use the letter to the Secretary as a way to say, okay, the committee has developed this tool, we appreciate your support in getting it disseminated, sharing it with partners. So that's another way to look at phase four is to have the letter and the report go together, which would certainly, in Pop Health, we've done in the past. It's just something to think about.

Rachel Seeger: Rebecca, you are correct. This is the model that we took with the beyond HIPAA work. So the subcommittee has a record in the past of developing a report and recommendations to the Secretary, and then that phase four that we had talked about in the presentation that Jacki did earlier would be to take these recommendations and findings and create a summary of report for the industry. And for consumers, for the public.

I think we called it a framework, and we don't have to get fixed on what we're going to call it right now, but it's the same strategy that we've used in the past on the subcommittee.

Jacki Monson: Okay, let's move into the plan. This is the first phase of the specific plan, we're going to look at the current landscape, evolving technologies, existing framework, AI, big data, and other expanding uses and users, and perhaps this is where we can enhance it to include public health; then internet of things, medical devices, applications, and emerging technologies; and then novel approaches to enhancing enterprise-wide security.

As we mentioned before, our approach is to have a virtual hearing, inviting lots of key stakeholders with various perspectives, and as we mentioned, we'll have lots of probably federal resources to help us with that, and then private sector, and then the goal would be to create a report of the environmental scan findings.

Phase two will be based on what we've learned in the environmental scan. We'll have the focus on narrowing the scope, developing models and illustrative future scenarios, laying out our assumptions and identifying areas of uncertainty, and then this will obviously reflect the work that we're going to take on about what specifically we're going to do. We're going to develop models and identify potential policy, practices, and technology solutions. So we might look at modifying that specific approach to clarify the technology aspect of that to make sure that it's neutral.

And then the deliverable would be to develop a draft framework, including benefits, levers, and relationships of policy alternatives. We're going to then prepare the recommendation for the Secretary, and all of the steps with respect to that will be framework, best practices, and potentially legislative mechanisms, if that's something that we identify through the hearing. Preparation of the letter, draft it, it will get approved by this committee, and then obviously, phase one, will be the report, primer and toolkit that specifically Rachel and Rebecca just spoke about.

This is our timeframe right now that we're proposing to focus on having a hearing yet this year, hopefully in the August-September timeframe. And then looking at to move through the other phases, through the rest of 2021, through possibly third quarter of 2022.

Then this is all the history that I think Rich was looking for, from 1996 on, when I think I might have been in elementary school, and it details specifically all the years and what the committee focused on and what we've delivered specific to privacy and security from there, which is a lot of really great work.

That really is the end of our scoping document. Is there any other questions, comments, or items for discussion?

Nick Coussoule: Jacki, I have one, and Rich has got his hand up as well.

The only thing, in phase I we're doing the environmental scan. I know we've talked about it a little bit, it's not necessarily explicit in the approach, is making sure we get the different perspectives as well as the different risks, as well from a consumer perspective, a healthcare practitioner perspective, maybe a researcher perspective, so just making sure that we've got the perspectives of some of the different entities that are involved in that explicitly as part of the panel. I know we did that in the beyond HIPAA work explicitly, I'm sure you've considered it as well. I just want to make sure that that's part of our considerations.

Jacki Monson: Yes, absolutely. We've noted that, and that's been something that as a subcommittee we've already discussed, so we will definitely try to get broad perspectives. I think it might actually help us that we're virtual for this hearing, to actually get more perspectives, because it won't require travel. So that's really what we're hoping for is to really get a good broad base of perspectives on this, and definitely talking about all the different perspectives, including consumers, which we've discussed this morning, as well.

Rich Landen: In phase 3, those last two bullets, a term I see missing, and wondering if that's by intention or not, is regulatory. The fourth bullet mentions specifically legislative mechanism, and then that third

bullet just talks about general levers, and like the ICAD report emphasized, was HHS levers from their program ability, like the Medicare program, conditions of participation, as opposed to a regulation under HIPAA or one of the other authorizing acts. So the basic question is why is regulatory not mentioned here? Is that off the table?

Melissa Goldstein: Absolutely not off the table. I promise you, we'll consider that, Rich. It's really just a word oversight, I think.

Rich Landen: I don't want to beat it to death, but phase 4, again, we're back here, prepare a report for healthcare industry and not the Secretary, so just please clean that up, and let the other -- make some consistency between the different sections.

Jacki Monson: We will do that.

Nick Coussoule: Any other comments, questions? Jacki, back to you.

Jacki Monson: That is all that we have today.

Nick Coussoule: Based on the robust discussion, I think you got a lot of really good support for this, in your work efforts. I guess I would just ask the committee members if you are good with our discussion, and positive what we're doing, please raise your hand, just so we can capture that for the record.

Rebecca Hines: Let the record note that all the members have their hands up except Margaret and Jim. Hands are up now, okay. So we have 100 percent. Let's move forward. Very good.

Nick Coussoule: Jacki and Melissa and Rachel, and others, thank you very much for a great presentation. I also thank the rest of the committee members for a very engaging discussion. You can always tell when a discussion is engaging, you get a lot of interest, so that's really helpful, I think, for us, and it'll be a great work effort for the committee to undertake.

Rebecca Hines: Because we have a lot of new members, I want to take a moment to review the way the process works. We have these full committee meetings during which subcommittees can bring forward work to all members. You can look at the schedule for the next year to organize your work keeping in mind when the full Committee meetings are scheduled, so that if you want to vote or need approval on something you've got it timed with your work and so forth. Just keep that in mind, and we will work with you to schedule the full Committee meetings after the one we have scheduled for September. But it's a good time, you have all the members here, just to touch base, here's where we are, and the next meeting will be the next opportunity to bring work forward to the full Committee. Many of you already are familiar with the process, but just sort of show that's how we plan our work out from Full Committee meeting to Full Committee meeting, with the subcommittee work obviously moving things forward in between.

Nick Coussoule: Great, Rebecca. I think we are ready to move on. We have our next session is scheduled for 11:30, with Dr. Jernigan. We have about 25 minutes between now and then. What I'd like to do is to go on a little bit to be able to maybe bring up and get a little bit of editing done on our 2021 report to Congress, so if we can bring that up and we can spend maybe the next 15 minutes or so going through a little bit more of that, that would be appreciated.

NCVHS 2021 Report to Congress

Rebecca Hines: Okay. I'm going to turn off the comments, so if you want to see the comments, look at your own copy. I believe we left off yesterday --

Nick Coussoule: We should be at section 2 --

Rebecca Hines: We were at the bottom of page 7, discussing social determinants of health, I believe.

Nick Coussoule: I had us just finishing up bullet one and just about to move on to bullet two.

Rebecca Hines: That is right. Here we are. Tammy, your hand is up. I didn't know if you still wanted it up.

Nick Coussoule: Denise, as well.

Denise Love: I will go first. I heard a couple things, and I've thought a lot about it. I think we have to be careful not to completely lose sight of the term social determinants of health, no matter how problematic that term is, because it's so pervasive right now. I just don't want to confuse the readers that there's something else. So, I think we had a solution yesterday; I just want to be mindful that we don't confuse the reader that might be reading about social determinants in one part of the industry and then have another term that they have to figure out.

Rebecca Hines: Denise, I would like to chime in that this is a report to Congress, and of course, HHS and others will be looking at it, and that is the term, I think Maya, or someone, put it in the chat yesterday. I do think we need to use the commonly used term, and I think it's great if we have a footnote that describes the limitations or the facets of the phrase social determinants of health that are problematic. But I would please encourage you to decide to use that term, because of who your audience is and because of the -- it's an accepted term now.

Denise Love: But I think we had some good points and texts that we could embed, not even in a footnote. I think we could talk about the limits of this term, but for the purposes of this report, we're going to use the common term, or something like that. I think we can wordsmith it. But I just wanted to weigh in on that before we got too far.

Lorraine Doo: Why would you want to not, especially after the presentation yesterday?

Rebecca Hines: Case in point. I see Denise Chrysler's hand.

Denise Chrysler: I was going to speak in support of Denise Love, and I fully believe social determinants should be used, that we shouldn't be swapping it out for other terms. Although we can always use qualifiers such as social determinants of health, those factors that affect health outcomes, whatever. But because it is so widely used and change in terminology happens over time - I did ask my administrative assistant to look for other terms that are being used, for example, Robert Wood Johnson Foundation is still using social determinants. Now she's looking for me for the pros and cons of that term. But I just want to speak in support of its use and possibly qualifying language where it fits.

Lorraine Doo: And WHO is using it. I sent out a link yesterday, as well as CDC, and that might be helpful to look at some of those references as well.

Rebecca Hines: Wu, your hand is up, and then Vickie.

Wu Xu: I just want to say I support continuing to use that term.

Vickie Mays: I wasn't here yesterday when you had the discussion, but just because I teach this stuff, just to give a sense, the social determinants of health we often talk about as a model or a framework. What you'll see within the healthcare setting is we often talk about social risk. So in some of the EHRs we're looking for social risk, which often are coming from the framework of the social determinants of health. Here, I think it really is the broader framework and that it's absolutely fine to leave it there, and one can make some qualifications. But right now you don't want not to use it, because it's embedded in the ways in which policies are going forth. That's different than practices.

Jamie Ferguson: I also wanted to express my support for what Denise said, and I think that since I'm the one who brought up some of the issues with this term yesterday, firstly, that social factors do not determine health, they're not deterministic; and also that people to whom these terms may be applied find them to be sometimes paternalistic or demeaning, and so there are issues with that particular term. At the same time, I recognize fully that it is now an accepted term. We should use it. But we should also, I think, explain, as Denise said, some of the issues and limitations with that terminology.

Melissa Goldstein: I think that if we explain what we want to say clearly enough in the report and footnote that why we've chosen not to use the term and include the term in the footnote, I think we'll be fine, because I think that the people that are going to read the report, both the public and Congress and congressional staffers, an everyone else, I think will understand when we say social facts that affect health. I think that people are very smart, and they will understand that we're talking about the same thing. We've just chosen a different way to talk about it.

The report itself is not centered on social factors. It's part of what we talk about. So I really think that as long as we are clear in our approach and with the wording that we use, and I think that everybody here knows what that wording is, and we have enough experts on the panel that that wording is fine, as long as we reference SDOH, I think we'll be fine.

Nick Coussoule: I think the question might be do we reference SDOH in the text and footnote the distinction, or do we reference different language in the text and footnote SDOH? Again, I think we're all conceptually in the same place, but it is a question of when we put the report together, how we reference the body of the report versus what we reference in the footnotes.

Rich Landen: Respectfully, I am going to have to disagree with Melissa. Absolutely if somebody takes the time to read what we write, that would be correct. It doesn't take a rocket scientist to understand what we're talking about. But since as was brought home by so many of the presentations yesterday, the term of art is SDOH, and if we fail to use that, then those readers who do not take the time to read in detail the language will not pick up on that we're talking about the same thing, if we use a different term. Since we have a 33-page report, it's unlikely that we'll get more than a relative handful of people who will read the details. So I'd be with the others who say we need to use the term SDOH but provide the explanation of why that is not a perfect term and, in fact, for the reasons Jamie spelled out yesterday and this morning, it's inaccurate in some senses.

Melissa Goldstein: I would say I don't think this is a creation of a new term, but I do think it's an opportunity for us to make the point that we've been discussing, which we all seem to agree with, that perhaps we should be more accurate in our terminology. It is an opportunity for us. I don't think we should create a new term, and I have a lot of confidence in the public.

Rebecca Hines: So are we ready to move on?

Nick Coussoule: Let's move on for now, Rebecca.

Rebecca Hines: I think we decided to lowercase value-based care from yesterday. Some other edits have been added. Any concerns with any of the edits that have been added on page 10?

Hearing none, there's a section 4 on the pandemic.

Wu Xu: Rebecca, do you want me to explain my idea?

Rebecca Hines: Please, thanks.

Wu Xu: Originally this statement are pretty generic that can be applied on any time that will have disparity among different population groups. So I think since we're talking under pandemics, so we should at least highlight one salient disparity happened during this pandemic. So I propose to add racial/ethnic and other population groups.

Vickie Mays: Can you just go back up? I wanted you to change a word at the level up. Can we change has exposed grave weaknesses to critical? And then now we're back to where we just were. Here, the word -- okay, it says increased reliance on the internet has raised awareness of the impact on health equity of broadband access. I think you want to talk about increased reliance on the use of technology to access COVID resources.

Rebecca Hines: So which sentence are you in, Vickie?

Vickie Mays: Okay. It's increased reliance on the internet. I just think it is increased reliance on the use of technology. Because there's other things. It's not just that. People don't have a computer, people don't have -- you know, they lost jobs, they lost their phone. So it just is I think technology.

Increased reliance on the use of technology has raised awareness of the impact -- I didn't think the impact was on health equity of broadband access, but the impact of the ability to access resources, such as broadband impact has highlighted health inequities. Because I didn't think it was a major differentiator between small rural and large urban communities, because we have these same problems in large urban communities.

Jamie Ferguson: I just wanted to add that I think yesterday we talked about the fact that broadband access actually is nearly ubiquitous. It's available to over 99 percent of Americans. The problem is affordability rather than technical access, and so I would just add the word affordable before broadband access.

Denise Love: Mine is easy. I think we're using grave, and do we want to put critical. We're in that paragraph, because we have critical in the title, and now we have grave down lower. The pandemic,

down, on my copy it has -- right here at the last sentence. Do we want to put critical, just to be consistent?

Nick Coussoule: I actually think that is a better word, as well, Denise. Let me read through this second sentence and make sure it makes sense consistently.

All right, Tammy?

Tammy Banks: I was just going to say can you -- we did already make a correction, affordable broadband access. Can we doublecheck how we changed it and just make it consistent before --

Rebecca Hines: It said affordable, but I can put a note in.

Tammy Banks: I think we had affordable and something else, accessible or I thought there was another word we had added, but just double check.

Rebecca Hines: You might be right. Let me just add in a comment.

So no changes until the middle of the next page.

Nick Coussoule: I think the grammatical ones are good.

Rebecca Hines: What about the text box, something about the addition of federal? In the absence of an effective hospital reporting system, COVID-19 reporting to local and state and federal, I think that's right. There's no doubt about that.

Nick Coussoule: Just from an editing perspective, you probably say local, state, and federal.

Rebecca Hines: Denise, I think this was your addition.

Denise Love: Yes. Yes, I took out that one part about state and local.

Rebecca Hines: Sorry for all the movement there. So now it says public health data now being captured vary in how demographic information is collected and released, limiting its utility especially in a public health emergency. Suppression of geographic specificity protects privacy, but does not promote understanding of risks and potential harms at the neighborhood, community, or population levels. That works.

Vickie Mays: I just want to ask a question of will we do any changes to this based on what we hear today?

Rebecca Hines: It is possible. We are meeting again on this and at least one more draft is going around to the full committee in between. So, yes. This report is on 2019 and 2020, but --

Denise Love: I hear what you are saying, but there may be some terms or wording that we can make more elegant, as we say.

Rebecca Hines: Yes. So actually, applauded was the word in the letter. We actually did not present that, so this is incorrect. We didn't present the WIN Framework. The Committee spurred its development and we applauded it when it was -- that's actually the word that the committee used, because for some

reason we didn't have consensus or something. My memory is a little vague. But I think that was the word that after much discussion everybody agreed on. So I think we need to use the word that was in the letter.

The WIN Framework, for those of you who are new members, is not a committee product, but the committee's work actually led to another organization developing the WIN Framework.

Anyone else? Okay, so continuing down, any input on the edits on these next paragraphs?

Nick Coussoule: Sorry, Rebecca, can you go back up just for a second? I think Vickie has her hand up as well. Just one question in the paragraph where it is to balance the importance of data access and use with appropriate privacy and security protections. Just the term appropriate, is that the right word? Should it be with statutory protections or require protection? I'm just trying to understand the distinction that we're making by adding the word appropriate. Maybe I'm the only one here.

Rebecca Hines: I think PCS added appropriate, but I'm not sure.

Nick Coussoule: Maya, you had your hand up and then Rich.

Maya Bernstein: In the past, we have tried to avoid use of the word balance because it applies when you have more of one thing, you have less of the other, and we try to use a different metaphor, rising tide lifts all boats, and so we can have more access and more privacy. We can have more privacy and security, and so forth.

Nick Coussoule: So was the question of whether the word balance is appropriate in there?

Maya Bernstein: Yes. That's the question.

Rebecca Hines: Rich, I was going to suggest the word consider, but do you have any suggestions?

Rich Landen: My hand is up for a different section.

Rebecca Hines: Okay, how about consider rather than balance?

Jamie Ferguson: My hand is up just to say consider.

Rebecca Hines: Great minds think alike.

Jamie Ferguson: And so do ours.

Nick Coussoule: I like the word, but I would think consider the importance of data access and use along with.

Maya Bernstein: So I would say if you promulgate a regulation, you're no longer considering it. You've passed a regulation. You might say promote or include or advance or some other thing that shows that you have already taken the action.

Denise Love: You could use weigh. I like balance, but you could use weigh the importance.

Maya Bernstein: Weigh is the same as balance for me, and in the past the committee has not used that word for the reasons that I said.

Jamie Ferguson: You want to address both things at the same time.

Maya Bernstein: That's exactly right.

Nick Coussoule: I think that appropriate or at least the concept is to make sure that they're both considered, right? So how the framing --

(Cross-talk.)

But my point is they are relatively competing activities, or at least seem to be competing activities, and our point is they both need to be thought of in the same vein and ideally it's not necessarily trading off, but you're trying to extend both.

Maya Bernstein: That is right. It's not supposed to be a tradeoff. You're supposed to try to accomplish both things simultaneously. There are ways to do that. I think in the past, it's true that you may -- society has thought of these things as a balance or a weigh, but that's not necessarily practiced. We have many tools now that don't require that.

Jamie Ferguson: Could I suggest the word address instead of balance or consider?

Denise Love: Consider is too passive in my mind.

Rebecca Hines: Can people live with address?

Nick Coussoule: Yeah, I like that.

Rebecca Hines: We have a time check. Dr. Dan Jernigan has joined us. So if it's all right, can we pause this until this afternoon, Nick?

Nick Coussoule: Yes, we can. Thank you all again for some work. So we'll move quickly onto the next section. We have Dr. Dan Jernigan joining us. He is the acting deputy director for public health science and surveillance at CDC, where he's responsible for strengthening CDC's scientific foundation, working across the Office of Science, Office of Laboratory Science and Safety, Center for Surveillance, Epidemiology, and Laboratory Services, and the National Center for Health Statistics. Primary focus of Dr. Jernigan's role is to provide leadership for CDC's data modernization initiative and to serve as an advisor for the CDC director.

Previously, Dr. Jernigan served as director of the Influenza Division of CDC's National Center for Immunization and Respiratory Diseases, where he was responsible for overseeing efforts to reduce the domestic and global burden of disease, and due to seasonal, zoonotic, and pandemic influenza. Dr. Jernigan received an undergraduate degree from Duke University, doctor of medicine from Baylor College of Medicine, and a master of public health from the University of Texas. He's board certified in internal medicine, has completed additional residency in preventive medicine. We are very happy and excited to have him with us today. So I will turn it over to Dr. Jernigan.

CDC Data Modernization Initiative

Dan Jernigan: This is an opportunity to talk through some of the things that we're doing with the data modernization initiative at CDC. I think many of you are familiar with it. Some of you may not be. I'm relatively new to this position. It's been two months now. Prior to that, as was just mentioned, was the director of the Influenza Division, but over the last year have been serving here as an incident commander for the COVID response or up in D.C. with the Joint Coordination Cell of the National Response Coordination Center for the COVID response.

So I come at this from a relatively program side of things, having worked on surveillance in a number of different informatic issues in the past, and served as board member for two years on HL7.

So what I'll do is -- let me know if this is the right thing -- but going to give you an overview of the data modernization activities, where we are right now, and some of the accomplishments that have occurred, and then we can have some discussion. Does that sound okay or correct? Excellent, all right.

So just for those of you that are not that familiar with it, the data modernization initiative is really in a series of a number of activities that have happened over years. I actually was involved in something in 1996 called the integration project, which was with Clare Broome. I don't know if you all remember Clare Broome from way back with the HIPAA and all that stuff that she was involved in, but it was an attempt to try to get at some common approach to data, and that's some accomplishments, but over years it's been a lot of initiatives but not a lot of resources to make it happens. The problem has gotten way more complex with way more technologies and a lot more systems that are out there, and it really came to a head with the COVID activity.

But this is a national effort to create modern integrated in real time public health data and surveillance that can protect us from any health threats. So it's driven a lot by infectious diseases, but it by no means is pathogen-specific.

So just a little bit about the why. You all should be pretty familiar with this next slide. There's been a lot happening over the years. This quote here was from Janet Hamilton, from the Council of State and Territorial Epidemiologists, prior to the pandemic, where she and I actually were on a government oversight or government committee hearing on surveillance and influenza, and really just she pointed out that everything is not working the way it should. Paper records, phone calls, spreadsheets, faxes, et cetera.

When COVID hit, it was astounding to see how many people were still using faxes. So very complicated electronic health record, vendors were having to figure out ways to format data so that it could be sent to a county or local health department in the format that they needed. So there's just significant changes that need to occur.

So the what. What is it that we've been doing with regard to this initiative of DMI? So I like I mentioned, lots of effort in the past, a lot of work Dr. Chesley Richards and Ileana Arias have been doing in terms of getting engagement with outside parties, working on very limited resources at the time, but putting together a plan and a capabilities that then could be implemented, and when the resources finally came, it made it much more capable for CDC to then move forward because of that work.

So in terms of public health data, this is for the this crowd something that I think all of us have been wanting to see for a long time, timely, granular, accurate, and accessible data, and just to the right there, the reality in the opportunity of things that in the past where we were reacting, we should be predicting, and that concept of predicting is something that comes out in the most recent of the

American Rescue Plan resources that we'll talk about in a little bit, moving from counting to understanding, from storing separately to sharing effectively, and this notion of how we're holding and making accessible our data is a significant component that we're working on at CDC.

Moving slowly to moving fast, and then using resources inefficiently and then of course connecting resources, and the cellphone there I think is a good indicator of how things are changing, and you all are seeing this in real time, that medicine is becoming distributed. People are accessing their medical care through their cellphone. They're talking with doctors that way, getting prescriptions, they are even diagnosing themselves now with Bluetooth-enabled cellphone diagnostics for COVID, and we know for flu that those are coming right behind that. How are we moving public health in order to capture the way people will be accessing medical care in the future? I think this is a big challenge for us.

So DMI, as it has arisen prior to resources, but certainly with the availability of resources, is definitely a unifying foundation. We've gotten considerable amount of partner support, both within the federal government and from a number of stakeholders outside the federal government. There's been congressional support with the funding that has come through the CARES Act and then now through the American Rescue Plan, and the public support, and then also it's been an interesting change of culture that we're beginning to see at CDC as well, that as groups at CDC begin to see how their own -- see how some programs are utilizing enterprise-wide solutions, they can actually see that and then be able to pivot and change the way that they've been doing things, because they're seeing those examples, and I think it's been eye-opening to see that getting things started, getting those kinds of pilot projects really helps everyone to be able to see how they themselves can participate in a more integrated approach.

So the funding so far, from the FY20 appropriations was $50 million, and in '21 another $50 million. So that's kind of a base fund that is provided to the deputy director for public health science and surveillance office that then gets distributed through CDC, and about half of it also goes out the door to states for various different projects.

That was the funding amount that started; of course, with COVID, the amount of funding going out to states is significantly increased through the CARES Act, but part of that CARES Act funding was also provided to CDC in the form of the $500 million that we are currently working through that allocation obligation and spend plan now in order to achieve interim successes for data modernization and also to set the stage for the infrastructure for doing the next wave of those kind of activities, as well.

So there are of course benefits to public health. It's empowering scientists. We hope to get better data, hope that we'll save time or encourage innovation. We'll promote collaboration and ensure sustainability. So these are benefits that we see we can get within CDC among those that are utilizing the different data systems and doing the programmatic work, but also for our public health partners at states, tribal local and territorial health departments, but also for other parts of the government, too, because the work that CDC does is not just for CDC, and I think you all would agree with that, I'm quite sure, and that we need to be making sure that what we're doing is going to have significant benefit to a broader audience and not just to CDC.

You should be familiar with the Driving Public Health in the Fast Lane, the 21st Century Data Superhighway document that CSTE and some others have put together. That was the basis for that congressional hearing Janet was at, in that it calls for five main areas. So if you want to know with DMI what's the tactical parts, what are you actually trying to accomplish, these five are part of that very specific implementation and improvement that's happening now; improvements in syndromic

surveillance and the national syndromic surveillance program, electronic case reporting, and we'll talk a little more about that, working directly by pulling those data from electronic health records, notifiable diseases through the traditional NNDSS activity and improving state surveillance platforms, like the NEDSS-based system. Electronic laboratory reporting, which is over 300 million of those messages have come through during COVID through the COVID ELR activity, and then vital records and vital statistics, those things that you all have a significant investment in.

So what are some of the priorities? Certainly, one is making sure that the data as it is -- as we're thinking about ways for that data to be collected and made available to folks to make sure that our partners are the ones that are going to benefit from that. Also making sure that parts of the USG and CDC are going to be benefiting from the greater access to that data, as well. So data collected being connected to CDC to our public health partners and then information being able to go back, bidirectional flow, through electronic orders and results or even in the form of making interpreted or clarified data available to our partners around the web.

Building a public health workforce. If you talk with anyone out there, the thing that really is the greatest need is developing a public health workforce of data scientists. There's just not enough of these people. They're not focused in their training necessarily on public health and getting them to be available for the needed changes is going to be a real challenge, but certainly one that we will be addressing with improvements in workforce. Then ongoing data modernization and other things underneath that comment that I can't see. Innovation.

So the how, what is it that we're actually doing, trying to lay a sustainable foundation. This concept of sustainable is absolutely critical, especially since the funds that are coming in are onetime funds and without a significant increase in a base, it's going to be very difficult for our public health partners to continue a culture or continue a cycle of development. So it's not enough for us to get a solution that just sits there and we're done. There has to be a continuous cycle of development incorporating new technologies, identifying new ways to do things that has to be an ongoing process.

So there are some focus areas, at least for me, that I've been focused on for the last two months, and that is making sure that we are building up, getting input on seeing, understanding, incorporating innovation. So that CDC has a long history of doing things in certain ways, and there is a need for us to be able to see how things can be done differently, but also to bring in people that can help us do that. So we are looking at ways to bring in fellows, bring in folks from the United States Digital Service from a number of different approaches to make sure that we can learn from them about what's happening in industry, what's happening in other parts of the government, et cetera.

We want to increase the amount of engagement that we have, and so I appreciate the opportunity to talk to you all, but in looking at our public health partners at the vendors, at the healthcare systems and healthcare providers and then also at the public to make sure that we understand what their needs are and that they have a role in helping us to figure out how to spend the money that we will be receiving.

Then finally the implementation needs to be sustainable, and so we will be working to take what are fairly high concept activities and initial projects and making sure that they are coordinated across CDC and across as much as we can the public health ecosystem so that we're getting to reusable enterprise-wide solutions, services that can be reused. All of those things that allow for something to be maintained more easily and to also be upgraded and innovated more easily.

So just to go a little deeper into that sustainable piece, I think this is a critical component, and if we end up with solutions that are not sustainable, we will have had another opportunity to get money and not have the long-lasting outcome from it. So this is a driving factor for a lot of the planning that's going on right now.

To be sustainable, components need to be reusable, secure, transparent, extendable, adaptable, and aligned. At the end of the slides, I'll talk a little bit about looking at DMI through the lens of data functions and making sure that we're looking to see that in each of the different kinds of functions of data through its lifecycle that we are making sure that we're trying to get the most reusable and most enterprise-wide approaches.

This is just a high-level diagram of the public health ecosystem, and for what we're trying to do, and that is to get better coordinated and seamless exchange of data. So there are a number of data providers and data flow, and that flow goes back and forth across this ecosystem. In a second I'll talk about some of the themes that we're using for bucketing the work that's happening, and so those themes generally fall into these data providers and core STLT-based platform improvements, and then theme two really looks at the CDC operated side of things in order to get the approaches at CDC to be modernized, and then theme three cuts across all of these in terms of trying to get to new kinds of standards by utilizing the U.S. data elements for interoperability as well as for the use of FHIR and other things like that.

So what we want to do is at that data provider level get solutions that are program-agnostic and pathogen-agnostic that collect as much data or data in as accessible a format as needed for as many different kind of uses of that data as possible, different programs, et cetera, at CDC.

So at NCHS there are some opportunities to try to get at the HCS survey that's done each year and doing that through the collection of information through EHRs and working with other parties at CDC that are trying to get information from EHRs, pulling that in, utilizing as few of these data hubs as possible for getting data around so that we do not have to have multiple different vendors in that space, and then working with the programs that have systems that are in place at state health departments and other health departments in making sure that they get modernized, and then finally getting the work at CDC, our own systems that get used by states, modernized as well.

How are we looking at this? There are several ways that we have made the projects understandable, and one way of doing that is through a roadmap or through what's really a high-level strategy. So this is one where an activity at CDC and with partners was done to identify what are the activities that we want to achieve, high-level activities, in a logic model, and what are the outcomes for short term, intermediate, and long term that we want to see?

So this is the high level of that. I am not going to go into each of these buckets, but I would direct you to the CDC website that actually goes through, and you can actually download this and then I'll make available some additional slides if you'd like that go through a lot of this activity here. So basically, we've identified what those activities are. We've identified short-term and intermediate and those very high-level long-term outcomes that we want to achieve, and we've mapped the different projects that are happening through the funding with the DMI to these activities and to the outcomes, and then we're using this as an evaluation tool for us to see are we are achieving those outcomes for those projects in the context of an overall strategy. So that's something that I think is really helpful for us to show that we're reaching some identified goals but also for communicating to legislators and others about the progress that we're making.

So the thematic areas for the DMI CARES Act funding, so this is for the $500 million that has come in, of which part of that has been allocated and obligated and set in motion with projects, and those have fallen into three main themes. The first is data sharing across the public health ecosystem. So getting data to all levels of public health with upgrades to core surveillance systems. So that's really that left side of that diagram.

Theme two is the CDC systems and service enhancements for ongoing data modernization, laying the foundation for cloud, machine learning, artificial intelligence, and advanced analytics at CDC. So this is getting to enterprise-wide systems, cloud capabilities, visualization and analysis tools that can be reused so that they're available for everyone to use so that you're not rebuilding all of those functions or approaches within a siloed system.

Finally, theme three are the new standards and approaches for public health data or public health reporting, adopting new standards and approaches that will give public health agencies access to richer data.

In terms of theme one, the key objectives were to develop and implement cloud-based approaches for automating data collection and supporting multidirectional data flows among STLT partners and CDC and USG, to reduce burden for data providers and public health agencies by automating it as much as possible, and ensuring systems and services are scalable, interoperable, and adaptable to evolving needs. This is turned into essentially the modernization of some select critical core surveillance systems, moving to program agnostic data sources from EHRs, emergency departments, and laboratories, and rapid extension of data lakes and services that support public health reporting and immunization data.

Some select accomplishments here. These are just examples of some of the things that have happened, but one of the biggest threats that we have and one of my other jobs right now is serving as the co-chair for the SARS-CoV-2 interagency group monitoring the variant emergence of the COVID viruses, some of which are having impact on our countermeasures like vaccines and therapeutics. So with that, we have been able to ramp up quickly to be able to handle the analysis and the ingest of over 800,000 unique SARS-CoV-2 genomes. That data is basically coming in, being able to be analyzed, and being able to be translated into different protein configurations and also to understand where certain epitopes on that virus are going to be likely to be impacted and therefore predict what the impact on certain countermeasures might be. So far, this data has led FDA to change its treatment recommendations late last week to essentially no longer recommending use of Bamlanivimab, a mono-antibody therapy.

In terms of vaccine response, over 140 million vaccination administration reports, average is 3.1 million records per day. This is a significant amount of data coming through and informing the response, and also allowing folks to see what's happening in terms of the amount of people being vaccinated. Lab reporting has increased significantly from thousands per week to millions per week, just for COVID-19. So this has been an exceptional lift that's happened, especially for the support from our National Center for Emerging Zoonotic Infectious Diseases.

In terms of electronic case reports, the standards were put in motion prior to COVID. The plan for that was in motion. The vocabulary, other things, have been decided for that. But with COVID, it's really given a push to move that forward. So now over 7,100 healthcare facilities in 47 states are sending electronic case records. They're doing that through working with the electronic health record vendors and others to incorporate this automatic case reporting. There's still a lot that needs to be done in that

space, but it's a glimmer of what things could look like if we had automated electronic case reporting for all reportable diseases and for other conditions, as well.

Environmental health monitoring, 2.7 million interactive maps and a number of environmental health measures that are available now for sharing and embedding, utilizing a common platform that other parts of CDC are now reusing.

Other things that we had significant increase to the amount of information coming through, the real-time emergency department reporting. Almost all states, a big majority of all of the emergency departments, and through this actually getting a pretty good completeness of race, 93 percent for the race variable alone. So this is something that can be used for not just COVID but for any encounter that's coming in through those emergency departments.

Working through the National Center for Immunization and Respiratory Diseases to set up scalable outbreak responses, essentially trying to make things better in NCIRD so that if we were to have another pandemic emerge, we would have systems that are utilizing databases and integration tools that could be easily scaled up for a major response.

Faster vital statistics. There's a reduced lag time, because of provisional death data that is now provided out, and others from NCHS can comment better on some of these things. Modernizing vital statistics across the United States, 67 percent of deaths reported now electronically within 10 days of death. Faster overdose death data, and expanded race and ethnicity data to inform the COVID-19 response. After the 2009 pandemic of H1N1, there was funding that was provided to help modernize some of the electronic death reporting, and it's really very encouraging to see now where we are with that, especially with the additional resources to make that really take off.

In terms of theme two, the CDC systems and service enhancements, the key objectives there to develop cloud-based enterprise-wide infrastructure, to enable data linking, sharing, analysis, and visualization, and include new and nontraditional data sources to complement existing systems. So far, we have had a significant impact with the implementation of enterprise-wide cloud services, utilizing Microsoft Azure platform as well as Amazon Web Services. Having both of those currently in place now to meet different needs and also as a continuity of operations approach.

And then modernization of legacy datasets to make data sharing faster and easier by not just lifting and shifting to the cloud, but changing the way that those databases are actually structured so that they can be more easily integrated across different programs.

New cloud-based solutions for cataloguing CDC's datasets, moving with Alation as a tool to do that. There's a whole world of library function there I think that we're beginning to do to understand our data better and then make it more accessible, and then new programs to train and hire, get the next workforce out there so that we can have folks that can do this kind of work in the field and at CDC.

I won't go through all of these, but there is increased data access, open data, we just made available, line listed deidentified data from all of the COVID cases that is available in different configurations on the website now so that you can use that information, use it for training purposes if you need to for machine learning, et cetera, but also for research and analysis. There's been significant improved visualization through a program that is our EDAV, our enterprise, data, analytics, and visualization services. Response solutions making it easier for us to get information at the border and for people that are needing to have contact tracing as they come through. Better connectivity by having cloud-based

solutions to ingest, analyze, and visualize the data. Enterprise services and data management. This is a very different approach for trying to have a single or as few ingress and egress portals for CDC so that when people are trying to share data with us and we're trying to share data out, we have as few places for that as possible.

Then the workforce. So CDC's Data Academy has delivered more than 1,000 hours of free training. Our data science upscaling program, there's new fellowships that are being started up, and there will be a considerable amount of effort here. Intertwining with efforts at the agency to also improve health equity metrics as well so that we have, we're addressing the issues of health equity with this as well.

So theme three is the one that is really just now getting off the ground and has a plan of the next several months with eventually at a HIMSS hackathon, having something in place with a FHIR pilot there. For this one, it is to develop and adopt new standards and approaches with real world testing of new standards for FHIR, notably for the U.S. Core Data Elements for Interoperability, standardized APIs, and electronic health record data elements, and hypothesis-driven discovery for fast-track implementation.

So far there is design and real-world testing of new FHIR-based approaches to interoperability that's just starting that's an activity that we probably doing with USDS and with a number of others to try to come up with the best way to test that through an acceleration, accelerator. Then alignment and advancement of existing standards, including the public health needs as prioritized in the U.S. data elements for interoperability.

Just briefly, in terms of what we're doing with USDS, there's one effort in particular called PRIME, or Pandemic-Ready Interoperability Modernization Effort.

The overall goal here is to get better, faster, complete, and accurate data to state and local public health departments so that they can take appropriate timing action. This was started last year as an effort to really help get the capture of laboratory data and the capture of the hospitalizations supply information and make that available for decision-making as quickly as possible.

This breaks down into three main components. USDS is working on data automation. So that's working on that data hub that sits at CDC where data can come in, be parsed to state health departments, can be brought into CDC, can be brought to other partners, as needed so that the transformation of that data occurs at the hub, rather than trying to do it at either end. And the use of an app that's been developed called SimpleReport for capturing laboratory test results and some other uses as well. The report stream is that data hub piece that I just talked about, and then finally improving the user interface of state systems by helping them to upgrade and modernize the way that they're doing that case investigations and other public health activities.

So getting near the end here. Just in terms of seeing DMI through the data lifecycle like I mentioned at the very front of the talk, we are as we approach the next tranche of resources looking at these data functions to help us to make sure that we're optimizing the development of enterprise-wide solutions, looking at data access and collection, data orchestration, routing, connecting, data hubs, et cetera, data hosting. Where is the data actually sitting and in what form is it being maintained and made accessible? Data cleansing and integration, especially with the use of cataloguing functions, analysis and visualization so that these tools can help for small events and for small needs, but also could be rapidly scaled up and made accessible as a common operating picture for a big response. Then finally, making sure that that data is disseminated and shared and made accessible for all those that need it.

One additional component here is the addition of prediction as well, that there is an increased focus right now on trying to get to better epidemic forecasting and outbreak analytics. So we hope that as we look through the functions of data and the lifecycle of it that we'll be harmonizing our efforts across the public health ecosystem to help us better collect, connect, track, and predict, allowing us to make more informed decisions and take more targeted action.

So just finally, there are of course a number of unmet needs and priorities, and we're not going to be able to do everything with the funding that's already been allocated to CDC. So we know that there are significant needs still for enhancing our partners' capabilities to conduct effective surveillance and response, lots that needs to happen with our own surveillance platforms that they're using to do their work in the states. We need to enhance CDC's capabilities to conduct effective surveillance and response, including the use of innovative nontraditional sources of data and improved forecasting of public health threats, and then finally, the workforce piece is just we have got to foster an environment of continuous learning and ongoing innovation with the right numbers of people and the right kinds of skilled people in order to do that.

So I think there may be one more slide. That's the last one. So I would direct you to our CDC website for the DMI initiative. There's a lot of different materials there, including these state-of-the-art lectures that were done so that people can not only see this high-level stuff but also can see very concrete examples that have occurred so far and can see those videos of folks describing what's happened at this point.

So I'm happy to take questions at this point. Sorry for a lot of talking, as well.

Nick Coussoule: Dr. Jernigan, thank you very much. That's an encouraging presentation and a daunting one at the same time. Frankly from my perspective, it's good to hear also that there are parallel paths between trying to figure out how to get better data faster, at the same trying to figure out how to make better use of what you have and when you get it. So it's encouraging that those are all also parallel paths. Sometimes we can get caught up in too much of we have to get it all perfect first, and then we can figure out what to do with it later, and then you're a lot further behind. So that's encouraging.

Let me call on the members. Wu?

Wu Xu: Hi, doctor. I am very excited to hear you present this modernization target. I'm a retired state health informatic official, so I'm sure when the CDC united as a whole agency and you try to modernize those core systems first, but you will work with the state because the data comes from the state most often. Then so my question is how can states leverage the CDC's modernization to modernize our systems? And just to focus on one, the state governance issue. So my question is when you make grants to the state, are you still using the traditional program-based grant mechanics or have a block grant on the public health information system modernization so in that way you can really force or encourage states has the overall whole agency approach like deal with this issue. So you can really increase interoperability, break down silo, at the state. So that's my question.

Dan Jernigan: To the earlier comment about it being daunting, yes. There are a lot of challenges there.

For the grant program, so the funding that's gone out through the CARES Act went out through the Epidemiology Laboratory Capacity grant approach, but the funding was made available utilizing mechanisms that are commonly used for the public health emergency program, the PHEP grants. So that activity then means it's more like a block grant.

So the funding that actually goes out, there's a lot of leeway with regard to how a state can utilize those funds. There will be additional funds that will be going out to states. It will likely go through that ELC grant mechanism again. But the hope and the plan is to make the guidance such that the implementation or the use of those funds will be achieving certain specific goals that have to do with improving state surveillance systems and other things that are very targeted.

The reason for doing that is so that the funding that doesn't go out and become an ambulance or a sidewalk. It is dedicated to the informatics approach, but exactly how that funding gets used to improve those state information systems is going to largely be up to the state I think to figure out how that funding gets spent. We will give guidance and some parameters around it. Several of the states use the NEDSS-based system, which is a federally supported activity, and so that one, that improvement will occur through other means, but the funding going out to the states specifically for data modernization, that will have guardrails around it.

Wu Xu: Just a comment. So the ELC grant will cover four of your five core systems, but not vital records. So I'm sure vital records will get some funding. I think that at the CDC level among the centers is a better you have same language as the interoperability and the interpretation between those systems in both grants. So because for me, I know the CARE grant so far is only ELC grantees, not vital records.

Dan Jernigan: That is actually a very important point that I can take, that as we're forming what that looks like, making sure that the program recipients are as broad as possible.

Wu Xu: Okay, so I want to make a comment. The HITECH funding in the 2009, they really required states have a uniform governance, they target governance. So that's really helped states have an overall governance mechanism entity to oversee all the interaction with the private sector. I think this time you need to have that have a requirement for states, cover all the public health, to make not program-based. Thank you.

Denise Chrysler: Thanks. That was a very exciting presentation, and at the same time, everybody notes daunting, and to see the kind of effort put into public health data. I spent most of my career as an attorney to a state health department, and so thinking about the role of law and that most data collection and surveillance and disclosure activities happen at a state level with not only variation among jurisdictions and law, but variation in law within a jurisdiction depending on the type of data, and how would you see funds for data, public health data modernization and efforts towards for this initiative being able to deal with the issues of variation of law?

Dan Jernigan: The issue of -- a lot of the stuff that we presented is really technical and technologic and information technology, but the really hard part is actually the policies. So you're absolutely correct. Is there specific funding -- are there specific projects for funding the policy? There are in that theme three component. So when we talked about developing the new FHIR standards and that approach, there is a fair amount of engagement in addressing policies there. That's really about policies to get to the use of those data and utilizing a newer approach with the new standards.

In terms of addressing some of the existing policy issues, I don't know that we have funding set aside specifically for that, but funding that goes to the states, I do believe, certainly could support that. But if you're asking is there directed funding in order to achieve those policy goals, I don't know that we have specifically, but I'd be happy to look into that and talk further about it.

Denise Chrysler: It wasn't so much funding, but the approach, and I know sometimes the approach is if you receive federal funds then there's functional standards, and there's conditions that help to align states or support of CSTE and their efforts to standardize case definitions, and all the various ways that variation in state law is addressed, but there's still so far to go.

Dan Jernigan: Right, so there is funding to some of those organizations for them to be doing a lot of that activity. We will be working across a number of different groups to try to develop an implementation plan to address this new, the DMI, American Rescue Plan activity. So that engagement will have to lay out what the policy challenges are and what those policy changes that need to occur, and those would be going through the traditional partners that we work with to try to address those.

Nick Coussoule: We have three more members with their hands up. I'd ask in the time sensitivity we want to be also cognizant of Dr. Jernigan's time. So if we could try to be a little bit quick, that would be helpful.

Vickie?

Vickie Mays: I will agree with my colleagues about the greatness of the presentation, but the question I want to ask is the priorities that CDC has established in terms of the modernization. One of the things is when Congress actually allocated this modernization, I know the Congressional Black Caucus and the Tri-Caucus were really very interested in the modernization in order to be able to get greater accuracy in terms of the collection of the data on race and ethnicity, and we're not there on that. I saw your slides, but I welcome you to come back to the presentation later and see how much missing data there is and the problem that's creating.

So I guess my questions are can you talk about the priorities you have in terms of working with states, because there are states that are not even collecting expanded race and ethnicity data, but yet you accept reports from them that then make the policy implications that you're making the decisions on be a little bit inadequate, because of that lack of data. So you can talk about the priorities working with states on that, the priorities in working with states on their immunization registries, which are missing a lot of race and ethnicity data, and whether there's any policies that you have that will make sure that you get complete data as opposed to accepting data and then it takes months to actually get those missing fields filled in.

Dan Jernigan: Yes, that is a complicated set of issues that we're all dealing with right now, and so as a part of the response to the vaccine taskforce and through other activities with the response, we have done sort of a landscape review of what's the problem here, where are we having trouble with the collection of that data is happening?

With that, I think they've identified, of course, that there are policy differences between the states, and so there is an effort, I think, in trying to get greater completeness of the race/ethnicity data to work with states in order to get those policies changed, because some of those are a part of the problem. Others are -- we actually did a USDS discovery sprint to try to understand better where some of the issues were and there were issues around where people are having to fill out forms themselves about not understanding exactly how to do those forms. There are systems that are -- there are times where it is not required to capture that data like in a vaccine encounter. A lot of this, they're looking at those in multiple different approaches to try to come up with a multipronged approach to see how to do that race/ethnicity completion.

By going to more automated collection, like electronic case reporting or through use of third-party lookups, where basically the data comes through and through privacy-preserving approaches, you ping it against data sources that can help fill in that information. There's a completeness that may be able to be done that way, but it's a near-term fix. What we really need to have is the availability of that data, the electronic health record, the ability for that to be captured through electronic case reporting or made available through other means, and then that being pulled and made available for a report.

I think that that's probably one of the better ways to do it, but it is a multipronged approach that is being worked on through the response right now with the White House.

Denise Love: Thank you. To follow on my committee members' previous comments, I don't think we have enough time to totally answer my question, but I will say that -- you mentioned the core datasets or platforms, but a lot of the information that's missing or the gaps in knowledge are coming -- are going to come from nontraditional data sources for the CDC, and having worked with hospital discharge data systems across states for 25, 30 years, I don't know, and all-payer claims databases, those are nontraditional. But I also think they will inform some of the gaps that the surveillance systems have shown to have, and so that integrative approach is something that I support, and I think it's probably implied in your slides, I'm assuming.

Dan Jernigan: There wasn't a lot spent on that, but it certainly the discovery of those datasets, the making available of those datasets in terms of access, assuring the quality and validity of that, and then the integration of that. So the integration of different datasets in order to fill in and get a better picture of what's happening, there's a science to that that I think also needs support as well. So there's components for research, where the plan is to have academics and have others figure out better ways where data can be integrated from multiple different data sources to help give a better picture of something.

Denise Love: Right. So I think some of that is going on at the state level. So this modernization and infrastructure may be unifying how that's being done, but one other question that you might not be able to answer, but I'm going to put it out there, is on your dissemination part I think what we realized during COVID is that there was really variability in public health transparency of the data, and so is there a mechanism for responsible appropriate and legal release of robust public health data to third-party trusted validators? So that we're all helping each other understand what's going on, because the politicization of COVID data was really distressing to me as a citizen, and I think it really undermined public health and it really hurts to see it play out in real time.

Dan Jernigan: I actually won't respond to that one. That one's very complicated.

Nick Coussoule: Last question then. Jamie?

Jamie Ferguson: Hi. First of all, I want to echo all the previous comments about what a wonderful presentation, and I want to build on a couple of the previous comments about the data that are collected, the data quality and data sources. Particularly within the past year, we have seen a need for increased standardization of data definitions and reporting formats. So within your theme three on standardization, I want to just urge you to consider the need for better standardization of data definitions for things that -- for example, health resources. Bed availability, ventilators, various different kinds of resources, where states and local authorities have just made up different reporting requirements that are sort of a jumble of incoherent variable data.

Similarly, we also need strong data definitions for sexual orientation and gender identity in electronic case reporting. That's something that's increasingly required in state law and it's not well standardized in terms of data definitions and formats. So I'd just urge you to look beyond the ONC's U.S. Core Data for Interoperability, because these things are not in that dataset.

Dan Jernigan: I agree with you, just the collection of that supply information and available resources, bed availability, all that is critical to the response, and so I think I hear the better standardization of that is, from you, and then the gender identity and those other variables, I think those have to be addressed sometime soon.

Nick Coussoule: I am going to call a time check here. Dr. Jernigan, thank you very much for your time and your information. I think the committee members would probably ask you questions for the rest of the day, but my guess is you have other things to do and we have other things on our agenda, as well.

Dan Jernigan: And may not have answers to most of them.

Nick Coussoule: But thank you very much for coming and the difficult part is we may ask you to come back again. So hopefully you'll consider that in the future.

Rebecca Hines: Also, Nick, I would like to use this opportunity to invite Dan, to call on the committee if there's a specific aspect that you are asked about. The committee deals with data standards and many of these issues. Obviously, interoperability is not as much in our wheelhouse, but please if there's a specific follow-up you'd like to make with the committee, we are the advisory committee to HHS, and we're here to serve.

Dan Jernigan: Great, and we really do -- we will have opportunities where we will be asking for help. So we'll definitely keep that in mind.

Nick Coussoule: Excellent. So to the rest of the committee members, we are going to be on break until 1 o'clock when we start our next panel. So please be on time for 1 o'clock, as we have lots of folks here to help us in our discussions. But we are now officially on break until 1. Thanks, everybody.

(Break.)

COVID-19: Capacity, Gaps & Quality in Collection of Race/Ethnicity Data – Expert Panel Session

Vickie Mays: Let me start by welcoming everybody and, in particular, thanking our panelists for taking time out of their schedule to assist us. I also want to, in particular, thank the leadership of ASPE. Sharon Arnold, who raised these issues that we are going to talk about in terms of anything the committee might be able to discuss. That would be useful, particularly for states and others.

So I am going to give a little background in terms of the panel which, as you can see, the topic is for us to really talk about COVID-19 in terms of the capacity gaps and quality in the collection of race and ethnicity data.

One of the things we know is that COVID-19 pandemic exposed a number of data gaps in our ability to be able to respond quickly and efficiently in determining the impact in racial and ethnic minority populations. So early in the COVID-19 epidemic, we had researchers, advocates, you know, states really

called out the lack of data identifying race and ethnicity in cases, in hospitalization and death, despite the fact that this data is part of our reporting requirements.

But it raised are issues about the methods of collection. What we saw was this issue about the lack of reporting in health data, electronic health records, and also in our mortality records. So the concerns that are raised are about the standards of incompleteness of health records, a lack of enforcement for completion in a timely manner.

We actually think about this lack of data quality due to incompleteness. And also the inability to link records I order to fill in missing data for national attention and leadership.

For those of you who follow this issue, what you will know is just yesterday, the Kaiser Family Foundation put out a brand new report on COVID-19 vaccinations and race and ethnicity. As of March 29th, 2021, CDC reported that race and ethnicity was known for just over half of people who had received at least one dose of the vaccine.

Among this group, nearly two-thirds were white, 9 percent Hispanic, 8 percent were black, 5 percent were Asian, 1 percent were American Indian or Alaskan-Natives. And even less than 1 percent were Native Hawaiians or other Pacific Islanders.

So the purpose of this expert panel today is to learn about the ways in which current practices and gaps in data collection, input sharing, linkages, and analytic methods could be improved to address some of the health inequities that we see in COVID-19. The lack of race and ethnicity data also impedes the capacity to identify outcomes that may actually be there in research treatment and helping us to identify evidence-based interventions at this sub-population level.

I am going to turn this over to my co-moderator, Denise Love, who will introduce the panel and add any additional words of welcome.

Denise Love: Thank you so much, all of the panelists, and thank you, Vickie, for opening this up. We have divided the panel into two parts. And the first part, panel A, will be talking about some of the inputs or the collection issue around race ethnicity data. And some of their learnings and findings that they have been working with during this pandemic.

And then the afternoon panel, or the part B, not afternoon, depending where you are, we will talk about some of the solutions and some of the possible workarounds and where we go from here. So this morning's panel, and I don't have the order, Vickie. I don't have it in front of me. But do you want me just to do the bios for each of the panelists?

So we are hearing from Dr. Betty Bekemeier, an endowed professor at the University of Washington School of Nursing who is a nationally recognized public health systems and practice space researcher, focusing on the structures and practices of state and local health departments regarding disparities and outcomes. The bios are available on the site and in the packets. So I am not going to read the whole bio here.

Then Vickie, will you take it from here because I am missing some of the bios?

Vickie Mays: No problem. Again, I think we have got on the website. So what I am going to do is introduce them by their titles. Trust me, they are great.

So we have Elizabeth Pathak who is the director of the COVKID Project. She is going to be talking about measuring disparities in COVID-19 in children and teens, the problem of missing race and ethnicity data.

We are also going to have Seth Spielman. He is the chief data strategy and analytics officer. And he is also associate professor of geography at the University of Colorado Boulder. And we will talk about uncertainty and demographic and social economic data, the use of differential privacy for disclosure control, and its potential impact on age and race ethnicity count.

Along with him will be David Van Riper who is the director of spatial analysis at Minnesota's Population Center. And in this panel, we will also have Alex Stokes, an assistant professor in global health BU School of Public Health. And will talk about race and ethnicity and the validity of cause of death assignment for COVID-19 in the US.

In panel B, which is actually the solution one, I will be talking about a predictive model, asking the question of what it is that we need besides race and ethnicity. We will have Glen Mays, who is no relation, we may look alike, but no relation. He is professor and chair of health systems management and policy at the University of Colorado. And we will be talking about collection and use of race and ethnicity data through cross-sector data-sharing network.

Michael Stoto to talk about framework to improve mortality, morbidity and disparities data for COVID-19 and other large-scale disasters which was a discussion that took place at the National Academies. He is a professor of health systems administration and population health at Georgetown.

And then Ninez Ponce will talk about the implications for COVID-19 equity and not disaggregating data in Native Hawaiian and Pacific Islanders. She is one of my colleagues at UCLA Center for Health Policy and Research and the UCLA Fielding School of Public Health.

I am going to be the timekeeper for the first panel. And I just want to say to my colleagues, particularly those of you in Colorado, and Dr. Ponce, who just lost her associate director of the Center, that we really appreciate your being with us and being able to do this work during a time where mortality is really hitting you in other ways. So thank you for being here.

So we are going to start with Dr. Pathak who is going to talk to us about kids.

Elizabeth Pathak: Good afternoon, everyone. My name is Beth Pathak, and I am going to be speaking today about the problem of missing race and ethnicity data for COVID-19. While my particular focus will be children and teens, most of the problems I will elaborate on are true for all ages.

I would like to thank the committee for the investigation to contribute today's panel. And I would also like to acknowledge my collaborators on a COVKID Project, Janelle Menard and Becky Garcia at WiiSE, and Jason Salemi at the University of South Florida.

I am going to move through my slides very quickly. So please do jot down slide numbers for any issues you would like to return to during the discussion. And I tried to make the slide numbers nice and big, so you can see them.

I am going to focus today on five types of surveillance data that we need to understand the epidemiology of COVID-19 in kids. And of course, these points apply equally to adults as well.

So to understand incidence of disease, we need testing and case data. To understand disease severity, we need hospitalization and intensive care admissions data. To understand disease fatality, we need death certificate data. To understand race and ethnicity disparities, we need all of these types of data.

So reporting of testing denominators is highly inconsistent across states. Rarely stratified by age, and there is only one state that reports tests for racial ethnic groups by age, and that state is Illinois. And I will share some data from the state of Illinois with you later.

I will spend most of my time today discussing the problem of missing race and ethnicity data in COVID-19 cases. This is the bedrock that we need to calculate disease incidence rates. All of these data originate with state health departments who share their data in two distinct ways.

They share directly with the public via their web dashboards, and they submit required data to the CDC. There is a great deal of mismatch between these two data sources. You would think they would be the same, but they are not. And I am going to return to that shortly in a lot more detail.

And an important point to note here on the side is that in the CDC case file, even though all states are reporting race and ethnicity. That is part of their requirement to report. It is missing for 45 percent of the cases.

Hospitalization data, we are in a much worse situation. As of this week, only 23 states report cumulative hospitalizations for children and teens on their dashboards. And none of these states report these data by race and ethnicity.

I will share with you a few slides with a very limited data available from the CDC in the CDC case files. And I will note that the DHSS hospitalization file contains no race and ethnicity data for any COVID-19 patients.

We do have a truly excellent, but unfortunately incomplete, data source for pediatric intensive care admissions of children 0 to 17 years old. This is a voluntary hospital registry which was initiated as a public service at the beginning of the pandemic by a pre-existing registry called the VPS registry.

And I really urge the committee and audience members to visit this dashboard as it is an exemplar of rich clinical surveillance data on our most severely ill children.

There are three primary sources of death data. State health department dashboards, the CDC CaseLine data file which contains a yes/no death indicator, and finally, the NCHS death certificate tabulations.

As you would expect, because of delays in reporting and processing, death counts from the first two sources are much higher than what is reported by NCHS. And I will just say here that I have been very impressed by and grateful for the speed with which NCHS has released a large number of COVID-19 death tabulation files.

 So now I am going to share results with you for all five of these types of data focused on the availability of the race and ethnicity data. In Illinois, for kids 0 to 19 years old, we can see in the chart on the left that there is a large difference in the extent of testing with much more testing being done in non-Hispanic white kids. And please note that these tests, these are not numbers of people who have been tested. These are numbers of tests that have been done in this age group. So we are seeing multiple

tests per person. It is not that 87.5 percent of white kids have been tested. But 87.5 tests per 100 kids in the population.

The test positivity rate is twice as high in Hispanics versus whites. And the incidence rate is 25 percent higher in Hispanics. So understanding testing denominators and test positivity rates are really essential for estimating true incidence rates. And we need better person-level data for all states, and obviously by race and ethnicity. So right now, this is a very big gap in our data.

Now, I am going to talk about the case data. In the COVKID Project, our case counts from weekly manual extraction of child and teen data from 50 state health department dashboards, plus the District of Columbia and New York City. Our counts are higher and more accurate than the reports you may have seen from the American Academy of Pediatrics because we correct for substantial non-reporting by New York state and Texas, two of the four largest states.

So here, we are interested in calendar year 2020. And we are comparing data from the COVKID tabulations and the CDC case file totals. And you will see that COVKID has more cases than CDC. But the situation is a little more complicated when we look at the state level.

So this chart shows the discrepancy between the COVKID case counts and the CDC CaseLine file. And this is for calendar year 2020 cases. So in this chart, both of the axes are on a logarithmic scale. So pay attention to those numbers that you see. It is on a log scale.

And the gray diagonal line represents perfect agreement between states. So when the states fall close to that line, the agreement is pretty close. And keep in mind that the CDC case file is for ages 0 to 19 years for all states. But the state dashboards vary in their reported age categories. Some report 0 to 19, some 0 to 18, some 0 to 17.

So the states that you see that fall well below the line, well below the gray diagonal line, are states which are not reporting all of their cases to the CDC. And in some cases, they reported very few of their cases to the CDC. So we know what they have reported on their dashboard, what they are submitting to CDC is just a fraction of that.

And committee members can find more details about these comparisons in the states that have problems in the slide supplement at the end of my talk. So at the end of my PowerPoint file, there are some supplemental slides there for committee members to review.

So our goal was to assemble the best estimate of cases. And we combined data from COVKID and the CDC case file to accomplish that goal. So as you can see, we use CDC totals for 13 states. An example is California because on the dashboard, California only reports 0 to 17. So we picked up a lot of cases by getting the 0 to 19 age group in the CDC file. But for 38 states, the COVKID total was better.

So after getting that best estimate of the total number of cases, we wanted to go back and impute race and ethnicity for those 45 percent of kids that were missing that information. So we wanted to impute only at the aggregate level to permit calculation of cumulative incidence rates.

Unfortunately, we cannot recommend imputation at the individual case level because there are few variables available in the file that allow anything beyond random imputation which we just don't think is good enough. Because there is a high degree of messiness, and there are not very many descriptive variables about the patients.

So at the aggregate level, our method for imputing is very simple. We simply assumed that the race and ethnic distribution of the missing cases was the same as the non-missing cases. And we did this at the state level. And there is some more information about the detail of the state-level imputation and the supplement.

Okay. In the interest of time, I am sharing the results for children aged 0 to 9 years only. And committee members can find the results for the 10- to 19-year-olds in the supplement.

On the left, you can see the raw data from the CDC file with 46.3 percent of cases missing race and ethnicity. After we corrected the total count, which increased by 160,000, so we see went from 643,000 to over 800,000 kids. And we imputed race and ethnicity. You can see the case distribution, the final case distribution, on the right. And remember that the imputation occurs at the state level. And at the state level, the race and ethnic distribution of cases varies widely.

And also the degree of missingness varies. So that is why the percentages don't match between the left pie and the right pie which are the national aggregates. And again, there are more details on this in the supplement.

The point of doing all this is to understand the impact of this missing data on incidence rates. So you can see in this chart that there was a huge impact of what we did on the incidence rates. The striped bars represent the uncorrected incidence rates. And the solid bars are corrected. And you can see how much the incidence rate increases for several of the race ethnic minority groups, particularly Hispanics.

And I want to note here that, to my knowledge, CDC folks have not performed any race ethnicity imputation when they published analyses of this CaseLine file. So they have simply excluded the cases that were missing. Race and ethnicity when they were looking at race and ethnicity data.

So now I am going to move onto the National Hospitalization Data. And so what you see here is data from the CDC CaseLine file, that hospitalization indicator. And keep in mind that this indicator was missing for 59 percent of the cases. And race and ethnicity is missing for 45 percent of the cases.

Now, we do think that probably missing may mean no for some states. But the hospitalization rates that you can see here are -- this is percent of cases that were hospitalized. 0.8 percent of whites, 1.5 percent of Hispanics, and 1.8 percent of the black patients. And the total hospitalizations reported for kids was 19,000 in this file.

Now, I am going to talk about the National Intensive Care Data. So this shows data from the VPS registry. And there were race and ethnicity data available for two-thirds of the patients which is over 1,800. And you can see the distribution with only a little bit more than a quarter of the patients were white. And nearly three-quarters were Hispanic, black, or other racial ethnic minorities.

And I would just like to emphasize that because this is a voluntary registry, we really don't know what the hospital non-participation rate is. So this is an undercount of the true number of kids who have been admitted to intensive care with COVID-19.

Finally, we have death data. And what I would like to invite you to do is to visit the COVKID disparities dashboard. We have quite a bit of detailed mortality data courtesy of the NCHS files that are submitted. And this is an example of one of our dashboards.

This is for children 1 to 14 years of age. We have the case distribution, mortality rates, and then the rate ratios with confidence limits. And you can see for this age group that both Hispanics and blacks have a death rate that is over twice as high as what we see for white children. And we also have data for infants, teens, and college-aged young adults on our disparities dashboard.

So now I would just like to summarize a few priorities that come out of this analysis. If you can just click through these as I read them. Race ethnicity data without age stratification is a very limited utility. Most state dashboards report race and ethnicity for COVID cases, but not stratified by age. And you really can't do much in comparing different racial ethnic groups if you don't also have age to either adjust for or stratify.

Missingness of race and ethnicity data needs to be systematically investigated and then actively intervened upon because it is still happening in 2021. So it is not like this is, oh, just an early in the pandemic problem, and it has all been fixed now.

It hasn't been. We still have testing cases and everything with a lot of missing information. And we need to pin down what the source is and come up with either the incentives or the consequences to get it to stop. I would like to see an end to excessive suppression of data, particularly death counts. We see a lot of that, and it really impacts the child reporting. And of course, we are predominantly suppressing deaths that are occurring among minority children.

Priorities for states include expand state dashboard reporting for kids, used standardized cut points and definitions to improve comparability among states. There is so much variety in what states are reporting, it makes it really hard to compare and contrast across geographic areas.

We need to include vaccination, testing, case, hospitalization, and death data. Report all data items by race and ethnicity. And we actually have a report card on the COVKID Project website that goes into little detail about grading states on the quality of the data that they are currently reporting.

We need mandatory hospital reporting of all COVID-19 hospitalizations of kids with demographics, length of stay, and discharge disposition. We need expanded clinical registries for all intensive care hospitalizations for kids. And we need emergency department visits registries.

This is something that folks may not have even thought about. But this is an important nexus for access to the health care system, especially for racial ethnic minority children. And it is something that we are not keeping track of at all, except for one exception which I will share with you in just a minute.

So now what I would like to do is share some information from the state of Florida with you to get you to imagine the possibilities. And this may come as a surprise to some of you because Florida is often in the news in a negative way. But I will tell you with great assurance that Florida has the absolute best and most transparent COVID-19 data releases in the country.

The state of Florida releases a COVID-19 CaseLine data file every single day of the week, seven days a week. And this has nearly 2 million cases released to the public on their website, no restrictions.

My colleague, Dr. Jason Salemi, who is an epidemiology professor at the University of South Florida publishes a daily dashboard with over 75 detailed metrics and maps and charts on every aspect of what is going on in Florida.

And the great thing about releasing this data to the public is that it permits the vast scientific expertise and resources that exist outside of government agencies to get to work on this problem. When we don't release data, then we are putting all of the burden on government scientists. And frankly, there simply aren't enough people and resources to do everything that we need to have done.

So this is just a screenshot from Dr. Salemi's dashboard. It is a little bit tiny print, but you can see that they actually release data in Florida on ED visits. And for patients of all ages, there have been over 125,000 emergency department visits for COVID-19.

And you can go in and select this by age, by county. There's a great deal of flexibility here. But of course, and I will just end on this note, the bitter irony here is that, as wonderful as this data resource is from Florida, the one thing that is missing is the race and ethnicity of the cases. So we will have a lot of work to day.

But I really encourage you to visit these dashboards that I have given you links for, so you can see what some of the possibilities are of what can be done. Thank you for your attention. Please feel free, anybody who is listening, to contact me with any questions. You have my email address there and also the COVKID Project website. Thank you.

Vickie Mays: Thank you very much. That was very informative. So let's move to Seth and David.

Seth Spielman: David and I are going to be talking about national sources of race and ethnicity data. So we are going to be taking about information that is available from the US Census Bureau for sort of the entire country. We are not going to be talking specifically about COVID related data, but we are going to be talking about the data that really informs denominators any time you want to calculate a rate.

And in particular, we are going to be talking about what we think of as the two primary sources of race and ethnicity data in the United States. The first is the American Community Survey. The American Community Survey, as it says in its title, it is a survey, not a census. So it is a sample of about three and a half million housing units in the United States each and every year.

The survey has about a 60 percent response rate. And because it is a survey, there is uncertainty associated with the counts that come out of the survey. So if the American Community Survey had a different sample and a different day, we would expect some changes in the counts. So there is some uncertainty in those estimates. And we are going to talk, in some detail, abut the magnitude of the uncertainty in the estimates of race and ethnicity.

The second source of information that is available from the Census Bureau is the Decennial Census. The ACS is sort of an annual snapshot of the country. The Decennial is a once every 10 years snapshot of the country. However, unlike the ACS, as we all know, the Decennial is, at least in theory, a complete count of the American population.

The key thing that we want to emphasize about the Decennial is that the published data from the Decennial have always, or at least in modern history, included some obfuscation of the data to protect privacy. So particularly for sort of smaller geographies like neighborhoods, there is swapping -- the data is directly manipulated by the Census Bureau to protect individual privacy.

And in 2020, the Census Bureau is adopting a new approach to privacy protection which is sometimes referred to as differential privacy or a formal approach to privacy. And that new approach involves

injecting noise directly into the data. And that injection of noise affects your ability to understand, in particular small populations. So racial and ethnic groups that are not a large part of the population in a particular place. We will talk about that in some detail.

So I am going to be showing you a series of graphs. Sorry about the X and Y axis. I will explain what those are. This particular graph that we are looking at here shows us the Hispanic population for almost every census track in the United States. We truncated 5 percent of the data just because there are some outliers.

So each blue dot on this graph is a census track. On the X axis, we have the Hispanic population. So the number of Hispanic people living in that census track.

On the Y axis, we have the margin of area associated with the estimate from the American Community Survey. So the red line represents the cases where the margin of error is 100 percent or more of the estimate. So that would mean the American Community Survey was telling us there were 100 Hispanic people in a neighborhood, plus or minus 150 or 110. The margin of era is actually greater than the estimate which means you know that there is somewhere between 0 and 100, 200, 500, whatever the number is.

The yellow line on the graph represents cases where the margin of error is 50 percent or more of the estimate. So any blue dot that is above that line represents a place where the American Community Survey is telling us that the Hispanic population in the neighborhood is 100 people plus or minus 50 or plus or minus 60. And that green line on the bottom represents census tracks where the margin of error is 10 percent or more of the estimate.

So when we look across all census tracts in the United States, for the Hispanic population, we see that 15 percent of all census tracts have a margin of error around the count that is greater than the estimate. So in 15 percent of all cases, the American Community Survey tells us something like there are 100 Hispanic people in this neighborhood, plus or minus 150. So we know there are somewhere between 0 and something.

In 50 percent of all census tracts in the United States, the American Community -- the margin era on the Hispanic population counts are 50 percent or more of the estimate. So there is still quite a lot of uncertainty for the Hispanic population at the neighborhood level when we look at the United States as a whole.

Key thing to notice is that many of the things that are sort of above the red line represent cases where the population count is small. As we move into -- as we look at neighborhoods that have larger Hispanic populations, in general, the amount of uncertainty as a percentage of the estimate tends to decrease. But it is still often fairly high. The margins of error are often over 10 percent of the estimate.

So here, we are looking at the exact same plot, same idea, same logic. Just that we are looking at the total population of the census tract. And what we see when we look at total population, is that very few census tracts in the United States have an extreme amount of uncertainty. Around the total population, very few have that error in that sort of medium category.

And about 35 percent of census tracts have a margin of error for the total population counts. That is 10 percent or more of the estimate. So there is still some uncertainty in the total population, but it is not nearly as much as when we start to look at subgroups of the population is the key takeaway.

I won't talk through all of these just because of time. But here, we are looking at the large sort of racial categories that are reported by the census. So if we look at the African-American population in the United States, about 25, 26 percent of all census tracts have a margin of error that is greater than the estimate which makes the data very hard to use because you know it is somewhere between 0 and something.

For Native Hawaiian Pacific Islanders, it is about 90 percent. For the Asian population, it is 40 percent. For the American Indian and Alaskan Native, it is about 80 percent of census tracts have a margin of error that is greater.

Now, census tracts are a very small geography. Census tracts are useful if you want to look at intra metropolitan variation. If we zoom out and we look at the quality of these estimates at the county scale, the data gets better.

So at the county scale, here we are looking at the Hispanic population using that same sort of type of graph. Only 3 percent of counties in the United States, when we are looking at the Hispanic population, have a margin of error that is greater than the estimate. And almost all of those are places where the Hispanic population is relatively small.

Ten percent of counties for the Hispanic population have a margin of error that is 50 percent or more. And 13 percent of counties have a margin of error that is 10 percent or more of the total population. If we think about the other ethnic groups, which are included in the notes for these slides, for African Americans, it is 12 percent of counties where the county estimate of the African American population from the ACS.

And 12 percent of all counties in the United States, the margin of error is more than the population count for Native Hawaiian Pacific Islanders, it is 80 percent of counties. For American Indian Alaska Natives. It is 30 percent of counties. So for lots of racial groups in the United States, even at the county scale, our estimates are highly uncertain.

I won't dwell on this slide too much. This is showing similar data. This is from a paper on this topic that we have published in demography. The key thing to realize here is that as you move into subpopulations, so if you look at the Hispanic population, the uncertainty is fairly large.

But if we were to look at sort of an age-specific group of the Hispanic population. Or in this case, an age and gender specific group of Hispanic population. The uncertainty around those denominator -- around those counts gets higher. Even at the county scale, we see which are these little diamond symbols.

Even at the county scale, the Y axis here, you can think of as the amount of uncertainty. Even at the county scale, we have a fair amount of uncertainty, if you wanted to determine say the number of Hispanic women who are over the age of 65. So across the denominator type data from the US Census, we see high levels of uncertainty.

There is this sort of Goldilocks problem where census tracts have very uncertain population counts. But they tend to be very small geographies. If we zoom out to counties, we often get good data. But if you are a health service provider or a state agency that wants to look at intra county variation for the purposes of program planning, all of that intra county variation is lost.

So there is really not a good sort of middle scale geography. New York City has created their geographies for reporting and analysis. We have developed software on a website to allow people to sort of create their own geographies. If we had more time, we could talk about that. We won't here.

And there are other methods to use to publish data from the ACS to refine estimates which again, because of time, he won't be able to talk about. But there are fixes, but there is a real problem with getting denominator data from the ACS in particular when you look at racial groups. That problem is compounded if you looked at age-specific or gender-specific racial groups. And with that, I will hand it over to David.

David Van Riper: Next slide, please. So now I am going to switch and talk about the Decennial 2020 census and talk a little bit about this new disclosure avoidance method the Census Bureau is going to implement. And look at the impacts that it is going to have on counts of sex and age groups and, by extension, sex by age, by race groups. And think a little bit about what that means for users in the long run.

So as Seth mentioned earlier, the Census Bureau used to use a technique called swapping which you see on the left here. And in swapping, they would target households and individuals who are particularly outliers. There was a couple of families in a census block that was Asian Pacific Islanders. And the remainder were all whites in that neighborhood. They would try to swap some of those households between census blocks to protect their privacy.

But it was a very opaque system. The policy decisions that went into that, the Census Bureau can't talk about because you can reverse engineer the swapping if you know what types of characteristics they are matching on and where they are doing their swapping.

So in the 2020 census, they are developing a differentially private method for doing disclosure avoidance. And that involves noise infusion. So in noise infusion, you are perturbing most of the counts to protect the privacy.

And the bureaus noise infusion method is designed such that the magnitude of the noise is completely independent of the cell count. Thus, we can observe situations like the tripling of the female who never attended school on the righthand side because of that noise injection.

This noise is drawn from particular statistical distributions whose shape is controlled by policy decisions made by the data producers. Policy decisions related to swapping are not publicly disclosed. But I want to talk a little bit about the policy decisions related to the differentially private algorithm because the one benefited the differentially private method is that all of those policy decisions can be made public.

This transparency is useful because we will know more about the disclosure avoidance. But on the flip side, those policy decisions can become political -- potentially political. And the way to think about these policy decisions is the bureau has to establish something called a private loss budget.

And they then spend that budget on various parameters. And the more money, the more they spend on a particular parameter, the more accurate the statistics will be for that geographic unit or table, the less they spend on that unit, the less accurate it is going to be.

So thus far, the Census Bureau has released the number of demonstration data products. And they have been targeting the geographic units you see here that I have circled in the gray boxes. So they have been focusing their spending of the privacy loss budget on states, counties, tracts.

Anything in the off bag that has not received an explicit allocation of the privacy loss budget. So we tend to see less accurate counts for those off diagonal units and more accurate counts down the middle.

In addition to the geography, they have to decide which statistics, which queries, do they want to make more accurate. So in this particular example from some early destination data, the bureau decided to make the voting age by Hispanic by race data the most accurate. And they had made the sex by age data the least accurate because they were privileging the redistricting, the political redistricting application of the data.

So now I want to show just a couple of results of what can happen when you see these counts come out. This is the census tract that I live in here in St. Paul. The red and blue boxes are the counts of individuals by sex and age for single years of age from the SF1 2010 data. The black dots are the counts of those sex by age groups from the October 2019 demonstration data products. And the green triangles are the same counts from the May 2020 demonstration dataset.

And if you look, for example, at the count of people under the ae of 1, you will see that we had about 40ish of those males and females in my tract in 2010. The October demonstration dataset showed almost no infants in my neighborhood. And the May dataset shows a more similar counts to what were published. But that doesn't always hold.

Even if you go up to a different set of age bins, so these are five-year age bins, we still see uncertainty in the counts comparing the red bars to the dots that are above it.

I just want to give a real quick real-world example for public health. What we are comparing here are asthma ED rates for Massachusetts towns. We are comparing rates of SF1 versus other denominators. You can just see the uncertainty here by the size of the place.

So the smaller the town is, we know very little about the age-adjusted rates. And as we make the sizes bigger, we get more consistency between the datasets.

These slides will be available. But just know that things are changing quickly. The Bureau is making lots of decisions about this. And we are going to learn more over the next six months. So you should all be paying very much attention to what the Bureau is doing over the next little bit. Thank you.

Vickie Mays: Thank you very much. Sorry we had to rush you there. But let's go to the next presenter. Dr. Betty Bekemier.

Betty Bekemeier: Thank you for having me. So my presentation will be a little different. It is less about the data themselves and more about local and rural public health perspective of the needs of public health practice system in rural areas in particular.

But when we talk about -- I want to situate this presentation, clarify that I will be talking mostly about governmental local public health departments. And that is because in this definition of what we call a functional local health department, they have pivotal positions in population health improvement and promoting health equity and reducing disparities.

And these local health department leaders, they are expected to use and contribute to the evidence base for public health. And we should be expecting them to provide expertise within their communities regarding promoting health equity.

They are often the hub, especially in rural areas for addressing health issues of equity. And the ones that communities look to for data, for information, for support around community decision-making and equity, and in particular, local public health emergencies.

So some of those, for me, started back in the 2008 massive budget downturn. And in Washington State, we were hearing horror stories of local health department leaders making really super painful budget cutting decisions like overnight.

So we did a study about how are these decisions being made. And the long and short of it is that all of our study participants talked about wanting to use data and evidence that they brought this up. But that the data were not available, they didn't have capacity to look at the data. They had to lay off their data guy, whoever their data guy was, who knows.

And there were dominant other forces, as shown in this figure, that were at play. But to a person, they all talked about the very small role that data made in decision-making and how much they wish they had better access to data. And that has not changed much since then.

I am going to distinguish a little bit between public health administrative data and community data around public health status. They are both critical to addressing health equity and particularly problems of rural areas.

But to try to compare for administrative data for public health agencies, they need to, they want to compare their agencies against others. And this is part of figuring out what are the best practices to measure public health impact on other services and resources.

And basically, it is not possible. No standard exists for the approximately 2,800 local health departments across the US for measuring their prevention activities or even expenditures. We have been working a lot on that. But we are far from it.

This lack of comparable data limits the extent to which policymakers, researchers, public health practitioners, can generate the evidence they need to make data-driven decisions. And that includes response to the COVID pandemic. And this is exacerbated dramatically in rural communities.

So around public health status, there is another project we have been working on we call SHARE-NW. But we focused on this project only on rural local health departments and in particular for the four northwest corner states. And many of these gaps and data access that you see that we have found here, that we have identified, are true for non-rural public health agencies. But they are particularly acute in rural areas.

More recently, we went back and interviewed rural public health leaders about data gaps during COVID. Here is what they said. The data, they had access to data, better access to data, but they were overwhelmed by the data volume. Public health practice leaders were. It was more than they could manage or handle.

They questioned the data trustworthiness. They had lack of capacity for assessing it, ensuring its quality for communicating with these data. And then they really lacked data for specific underserved groups. So it was all compounded for them during COVID.

So over the years as our team, the Public Health Activities and Services Tracking, or PHAST, team, has worked on data issues and related research for health departments, we have developed this model. Now, this model really represents the needs of governmental public health practice people and depicts this kind of virtuous cycle of factors necessary for integrating standardized administrative data into public health systems, as well as supporting deeper data use and access.

So these three different red-headed buckets here are all very interrelated. And all need to be attended to, to really address some of these gaps. That has really been found for us in our work.

For public health practitioners, they really have a need for access to relevant data. And it is often expressed as a need for those data for planning, for decision making. And access doesn't mean like giving a bunch of spreadsheets to them. That is not access. They can't use that.

They would also like better to be able to better use the data they have. And they talk about their need, their challenges, their potential use of data are all critical to designing effective strategies for community health improvement.

So through our work with public health practitioners over the years, they have really described old data, reporting systems, lack of centralized systems. And when the data are collected, it is often done by several different agencies or state public health divisions. So they have to get data from numerous different places. They question its quality.

And when they want to answer newer oppressing problems like with the opioid epidemic, now COVID, et cetera, the data have often not existed to even lend themselves to understanding the degree of the problem and looking at differences among marginalized populations. They have also talked about the lack of expertise in their agencies for using data.

And then in terms of understanding how data are used and how they want to use them, they always talk about wanting to be able to use data to better communicate within their communities, with their boards of health, et cetera, with elected officials. And for discovering where the needs are for specific populations.

So part of this is on us as researchers too and the generation of data. And our primary activity has been to sort of standardize data measurement systems and collect relevant data, make use of them and present them back to and work with our practice partners around making these data more useful for generating evidence, et cetera.

So we have worked with a lot of existing administrative data with health departments and done a lot with that and done some important research that analyzed data that is with questions driven from practice. But it is a tremendously huge task. It is no small task. There are great limitations to these data, et cetera. So it really underscores the sort of iterative nature. You have to work with practice. But it also underscores the need for standardized data.

So this last piece of the model is really about bridging this gap and really highlights data access. It explains that you have to provide a bridge between research and practice that really provides these channels to make it accessible and usable, et cetera.

We have looked at, and are working on, things like centralized data reporting, user-centered designed data visualization to make the data really, truly accessible to practice people. Providing that training and education to build capacity, et cetera. And all of these channels need to be effectively attended to.

So with some of our current funding, on the left are those data gaps. And with our current funding, we are trying to address some of these more immediate solutions, at least in rural areas in our region. And addressing some of these gaps. I won't go through them all.

But you will note that a lot of the sort of low-hanging fruit here is in the form of the capacity building and providing training and helping folks make better use of existing data. And we have been doing a ton of data linkage of data from many, many different sources to make it more useful to people on the ground in their local decision-making.

But finally, in terms of sort of larger, broader solutions, our lessons learned regarding overall improving data used specifically for rural public health is that sort of on the administrative side, better data are really critical for more equitable resource allocation, for really sort of being able to see where researchers are going and how they should be allocated.

And that requires data standardization. Such data standardization requires substantial resources and time, coordination. It frankly needs to be mandated. There need to be incentives involved, et cetera. We have been working a lot around that.

In terms of public health status data, data need to be accessible for people who are doing the work in these local communities especially in rural areas. And just providing Excel spreadsheets is not accessibility for them. There needs to also be training and using the data, collecting more data, sharing and engaging around those data with their communities. And then resources for over sampling and data collection.

And Dr. Pathak mentioned it too earlier that the whole issue of suppressed data for small numbers in rural communities is a big issue for them. And that kind of gets to the oversampling piece or the data collection, rural public health officials often only see asterisks where the small numbers of data are suppressed.

So when they want to explore variation by race and ethnicity, their hands are often tied around that. And even sometimes in terms of more general data. So we need to find a solution for this to at least for somehow making these data more available to public health officials.

And here are references on these slides. And I am all done.

Vickie Mays: Thank you very much. It was very informative.

The last presenter in this part A is Andrew Stokes. And Andrew, do you want to get started?

Andrew Stokes: Thank you Dr. Mays for the opportunity. I am really pleased to be here. I am very excited to present some of our findings to this group, this distinguished group. And thank you for the presenters who have gone so far.

I want to preface this by just giving the bottom line of our findings because I think with endless Zoom calls all day long, we all tend to lose focus over time. So let me just start by saying that our team is looking at small area patterns of excess mortality in the United States across counties.

And we are finding that the racial and ethnic disparities in COVID mortality are significantly underestimated when we rely on the official NCHS tallies. When we use our various methods and tools for incorporating excess mortality, we are seeing much larger disparities for Latinx individuals, for non-Hispanic black individuals, and for native American individuals.

And so there is an imperative here to get these data out there as soon as possible to really demonstrate that we are not getting to the full extent of the mortality crisis that is unfolding in the way that it is affecting different groups differently. So let's move on to the second slide.

I will talk about different types of excess death estimators. How many excess deaths have occurred? What is known about racial disparities? I will talk a little bit about our forthcoming findings which are coming out in PLoS Medicine and in several other journals.

And then I want to talk about why there are race and ethnic disparities in underreporting. Like why are the death certificate data not getting it right for racial ethnic groups, and what are those mechanisms of underreporting? And then I want to talk a little bit about future directions for research.

So excess deaths are defined as the difference between the observed number of deaths in a specific time period and expected number of deaths in the same time period. There are a lot of estimates now of excess mortality at the national and state level. And what I think we need to really do as a next step is to create these estimates at the county level, so that we can really drill down to what is happening across local areas of the country.

So in thinking about excess mortality, we have directly assigned deaths. Then we have excess deaths that were not assigned to COVID-19. Many of these deaths are misclassified to other causes of death such as circulatory diseases, Alzheimer's disease and related dementias, to an unspecified garbage code category and to other causes.

There are also indirect deaths that are related to interruptions in health care services. Interruptions in the provision of acute health care, emergency department visits, as well as these broad social and economic consequences of the pandemic, effects on social isolation and loneliness, and so on.

So we know based on NCHS estimates that as of 3/20, we had approximately 526,000 deaths. An unacceptable number of deaths to do COVID as a result of the direct effects of the pandemic. So these are the direct official tallies as of 3/24.

We also know that there are clear patterns in misattribution of deaths to other causes. We can see that there are excess ADRD deaths. There are excess diabetes deaths. So we have some sense of where these deaths are ending up.

One thing that we desperately need to explore is the kind of mechanisms that explain why these deaths are ending up in these other categories. And we ought to spend more time looking at these garbage codes which might, in some cases, indicate, to be frank, foul play. Sometimes, deaths may be misclassified due to political motivations. And we need to uncover those and to get to the bottom of the story here for excess mortality across local areas.

So we also know that there are these indirect deaths. Emergency department visits have declined significantly, especially early in the pandemic. This is a MMWR report that I think probably everyone saw. And we can see that these declines occurred for many serious reasons related to acute care.

So what do we know at the national and state levels? Based on kind of the best evidence we have so far from the NCHS, from Weinberger, from Wolfe, and others, we gathered that about 22 percent of all excess deaths were not assigned to COVID according to Weinberger.

If we look at NCHS-based estimates as of 3/30, they present a range of excess deaths between 538 and 655. And the percent of excess deaths not assigned to COVID, they estimate as somewhere around 20 percent which is fairly consistent with Weinberger.

So what is known about inequities in COVID-19 deaths? A lot is known about inequities in the official tallies. We have great evidence about this. So here is data from the NCHS on the percentage of COVID deaths by race and Hispanic origin.

We know, as others have discussed in this panel, and has been widely discussed among experts in race and ethnic differences, it is very misleading to look at crude death rates when it comes to race and ethnic differences. We have to age standardize.

And as you can see here, once we age standardize or stratify by age, you see that really alarming differences emerge for Hispanic and Latinx individuals, for non-Hispanic black, non-Hispanic American Indian or Alaska Native, and non-Hispanic Native Hawaiian or other Pacific Islanders. So there is really problematic differences here that we must not get lost in the narrative.

So here is a preprint that we have online that is forthcoming in PLoS Medicine. And what we want to try to add to the conversation is this kind of drilling down to the county level. Because there is just only so much we can learn from state level data.

And if you look at the other papers on excess deaths, really very few of them have been able to look at kind of social structural factors that are related to excess mortality. Because at the state level, you just don't have the detail. There are some exceptions.

There is, for example, some wonderful work being done in California by Kirsten Bibbins-Domingo and Maria Glymour and that team that has been looking by race and ethnicity and occupation. There is great work by Elizabeth Wrigley-Field and colleagues in Minnesota. But by and large, most of the excess mortality estimates do not give us a lot of information on the determinants and patterns. So that is what we hope to add to this conversation.

So we ask, what are the total number of excess deaths. We also ask, how does the percentage of excess deaths not assigned to COVID-19 differ by county level factors? So we are really interested in these kind of mechanisms of underreporting.

I am going to skip most of this. But we take advantage of the spatial variation and COVID-19 mortality across US counties to estimate its relationship with all cause mortality. So we are looking for signal, this imprint of COVID-19 and the impact that it is having on patterns of all-cause mortality across counties.

We use NCHS provisional county level data from January 1st to December 31st. We allow for reporting through March 12th in order to build in lags and death certification. We also construct a historical comparison period using data back to 2013 from CDC Wonder. Our county level covariates come from the RWJ County Health Rankings Project and other sources.

We have US Census Population estimates for historical years as well as by special request for more recent periods. And we explore model stratification by quintiles of county level factors in order to look at the social and structural factors associated with under reporting.

So here are the counties in our dataset. In this current paper, we use a dataset consisting of about 2,000 counties which are counties with 20 or more direct COVID deaths as of March 12th, but for the period of 2020. So you can see we still have some gaps in these data that we are able to fill as we go forward with more recent data.

So here is the basic relationship between direct COVID and all-cause death rates in 2020. There would be no excess mortality over this period if all the observations fell on the black line. But as we can see, in aggregate, the relationship is off that line. It is higher. The slope is steeper which suggests that there is this gap.

There is this discrepancy and that there are excess deaths occurring which is not surprising. And when you stratify that by region, we are seeing a great deal of underreporting in the south and the west. Those seem to be particularly problematic with respect to underreporting.

Overall, in aggregate, for every 100 deaths assigned to COVID, the number of all-cause deaths rose by 120 suggesting that 17 percent of all excess deaths were not directly assigned to COVID-19 on the death certificate. So this estimate is very much in line with Weinberger and other NCHS estimates based on other methods.

So we are reassured that our approach is kind of replicating findings of other studies. And we can break it down 367 directly assigned deaths over 2020, 74,000 of these unassigned deaths, as well as leading to a total of 441,000 excess deaths in 2020. And we have confidence intervals around these estimates.

So to get to the bottom line, if you look at non-Hispanic black, and you just look at the blue line. The blue line is the observed direct COVID-19 death rate, you can see that it appears that black communities have lower COVID rates than white communities based on crude mortality rates. But the minute that you factor in the hidden deaths or these uncounted deaths, very stark patterns emerge with higher disparities between black and white communities.

I want to end this presentation with some mechanisms. So why are black deaths and deaths among minoritized populations being differentially underreported? It could be differential access to testing and health care. It may make it less likely that residents receive a COVID-19 diagnosis before they die.

It could be a higher prevalence of comorbidities due to racial inequities in health. It may make accurate cause of death assignment less likely. It could be more indirect deaths occurring in black and brown communities due to the social and economic consequences of the pandemic.

It could also be bias in death investigation systems. The death investigation system in the US is irregular with some states relying on statement medical examiners, while some counties rely on county-based medical examiners, coroners, sheriff coroners or justices of the peace. Coroners, in particular, are often elected officials with little or no medical training.

And furthermore, there is potential for interpersonal bias and political attitudes about the pandemic which could influence whether COVID-19 is included on a death record or whether post-humus COVID-19 testing occurs. We are very concerned that, in some areas, coroners have really come out saying COVID is a hoax. And we really have to wonder whether they are willing to assign COVID to the death certificate.

The final slide here is home deaths, areas that have more underreporting of COVID-19 deaths could represent counties where a greater percentage of deaths are occurring at home. Dying at home could increase the likelihood that a death certifier or coroner would be involved in the death investigation process and decrease the likelihood that COVID-19 testing would occur.

Thank you for the time.

Denise Love: Thank you. Vickie, should I take it from here? We may leak over a little bit into our break because I am fascinated. This is an excellent panel.

Panel A, as we indicated, identified some of the gaps and challenges we have, and COVID has exposed those. And what we hope to accomplish with this next panel, panel B, are practical solutions to maximize what we do have to gain insights that we need now.

I wanted to add one panelist that got left out of the introductions is Dr. Mark Cullen, the founding director of Stanford Center for Population Health Science has retired.

We will start with Vickie Mays from NCVHS, then go with Glen Mays, Mike Stoto, Ninez Ponce, and Dr. Cullen.

So Vickie, I am just going to turn it over to you right now to open up.

Vickie Mays: Thank you. What I am going to try and do is to really focus more on the conceptual part and less on the results. The reason that I started this work is that when the epidemic first began, it was the issue of pre-existing health conditions. And many racial ethnic minorities who were obese or had health conditions thought they were being blamed.

And we knew that one of the things that everyone was asked to do is to figure out an equity metric. The biggest issue that everyone is struggling with right now is what is the best equity metric to use.

So part of what we did was took LA County which, of course, has a very large and diverse population and has lots of different subpopulations. I worked with the UCLA Center for Neighborhood Knowledge. My colleague, Paul Ung, who is an urban planner and an economist by training.

What we did was decide what are the things that put people at risk. So the big issue for us that we think is the way going forward is to think about vulnerability and risk to becoming infected as opposed to just kind of a spatial orientation of this is where cases are. So what we did was we developed this medical

vulnerability indicators which have these four things that we though were very interesting in terms of being able to capture kind of social risks.

So here, when we did pre-existing health vulnerability, part of what we did was we used -- at the time, this was actually done back in November. So we used the available scientific literature to say what are the health conditions that seem to be associated with coming down with COVID? And then we also checked with some of our health experts who were working in the emergency room and other places where they were actually seeing the cases in case things were not getting to the scientific literature fast enough.

In addition to actual health issues, we also included health status. We included mental health. And we included food insecurity because the pulse survey was telling us food insecurity was a very significant factor for people who had COVID-19.

We did barriers to accessing service as one of our equity modules. But it isn't what you usually think of as barriers to accessing service. Because it wasn't just about health services. But we were looking at lack of broadband access. We looked at whether or not people had a vehicle.

LA is huge, and so if you are getting on public transportation, you are going to be more exposed than if you are in your own vehicle. We looked at limited English proficiency, lack of health insurance, and citizenship.

This was also probably, if I had to say, what works best of all things. And it has a lot to do with the built environment. The density of the population, the extent to which people are on top of and close to each other.

The extent to which the building structure in communities, are housing really close to each other. And what is the density of that housing structure? Are these apartment buildings, housing projects, et cetera, with lots of people?

And then we looked at the in-unit housing crowding because in LA County, as in most of California, rent and mortgages are very expensive. And so you will see more people living together.

But we also have racial and ethnic populations that have three generational households. And so again, this issue of being able to do the public health advisories are sometimes not as easy as people think. We use the very interesting dataset called the California Parks and Recs dataset in which we looked at the availability of open space, the extent to which people can go out to walk, to be able to do that and not be on top of each other.

Then, of course, as required by CDC and many others, we had the social vulnerability index. We actually used CDC's social vulnerability index, but what we did was that it was only available at the census track level. Our data, for example, was available at a ZCTA level so we adapted it to be able to move the data that we had.

And here, what we are looking at is social economic status, the household composition, minority status and language, housing type and transportation. So you can see that in essence, we created a model that has to do with vulnerability to exposure. And then that exposure being the extent to which that person is likely to become infected.

I am just going to show you one of the outcomes of it. Here is mapping. So we mapped an LA county. What areas have the highest vulnerability for exposure and hence becoming infected. And a mapping has been used and is being used by a variety of entities within the county to try and understand where should we be putting resources. So where should testing be, where should that see distribution being rolled out?

Here, what you can see is we can give you space using the variables we have. But we could also do population. So for preexisting health vulnerabilities, what you can see is which of the populations are most likely to have that?

So in essence, if what you do is really combined the notion of the racial ethnic population, and then the area of vulnerability, what is you are getting is the ability to really target quite well where to go, where to roll out resources. So I am going to just quickly show you the other maps, so that you can see them.

This is barriers to accessing services. So if we are going to do a fix, we can determine what the fix needs to be by area. We can determine who that fix should be. This is our built environment. What you notice is the map now looks very different than it did in the beginning. So you can see that built environment is not an issue for everyone because the housing will vary significantly by income.

And then again, you can see the differences here in terms of who this is likely to be an issue for. And then let's just do the last one on CDC's adaptive social vulnerability. You see our map again begins to change. And you begin to see who is similar and who is different. So this really allows us to be able to roll resources out very specifically to particular groups and particular places.

So I am going to stop there, Denise.

Denise Love: I wish we had all afternoon. And so now, we are going to move right on, I am sorry, to Dr. Mays for 10 minutes. Then we will just keep rolling and then have discussion at the end. I hope we have time.

Glen Mays: Fantastic. Thanks so much for letting me join this panel of really great discussion, very important issues here. I am going to focus in on another potential solution that has come up in some parts of other people's talks today.

And that is thinking about community networks as a potential solution. Networks of community-based organizations working together to resolve health and social issues for their communities as potential solution, both in collecting better data on race and ethnicity, but also in using that information to actually reduce health risks associated with COVID-19 and other large-scale health events.

Much of what I will be sharing with you today is data information that comes from the National Health Security Preparedness Index. It allows us to measure the strength of preparedness activities that exist across the country and states and communities, understanding how that varies over time.

In particular, I will share with you some of these data that help us understand the importance of these multi-sector community networks and coalitions. The relationships that exist across community organizations, working in health care, public health, and social services. And what we know about what role those networks play and actually responding to -- preparing for and responding to large scale health risks like COVID, but also other health risks.

And then importantly, what the implications are for having better data on race and ethnicity to empower those networks to do what they do best and helping to protect communities from these kinds of health risks.

So one of the reasons why we might want to focus on networks is we think about the kinds of activities that we use to prepare for and respond to large-scale health threats like COVID-19. These are many different kinds of capabilities that are needed at the community level.

And no single organization can do all of these things in public health, in medical care delivery and social services. And so we really need to think about collective actions. And networks are a way of supporting collective actions, helping to facilitate communication, coordination across organizations at the community level.

The health security index is an index that was launched now about eight years ago originally by CDC to measure the strength of these protections capabilities that are needed to prepare for and respond to large scale health events. The early work on this index was conducted under the auspices of the CDC for the last five years or so. This index has bee produced and refined with finding from the Robert Wood Johnson Foundation.

Like any index, it aggregates lots of data and metrics from imperfect data sources, a total of about 130 measures go into this index, pulling from about 60 different national data sources. And we produce the index annually every year for national estimates, state estimates, and then county-level estimates.

This slide just shows you the domains and sub-domains that are measured in the index, and domains around surveillance, planning and engagement, incident management, health care delivery, countermeasure management, and then environmental occupational health. The domain I will really focus in is that community planning and engagement domain, and specifically our measures of cross-sector community collaboration. That is the place where the strength of the network is really measured in this index.

So we can use the index now. It has been produced since 2013. So we can look at trends over time at a national level. These kind of preparedness activities have been improving over time. And in the average US community, both overall and in the specific domains.

But we also see quite a bit of variation across the country. And again, how prepared states and communities are in terms of having these capabilities in place.

We also see, not surprisingly, wide geographic variation in the strength of these protections that are measured in the index over time. And we can see kind of regional clusters of communities that are stronger and weaker on these kinds of protections.

So one of the things that we do now that we have carried out over the last year is to look at how useful the preparedness index metrics are in actually explaining the variation that we see in COVID incidence and in COVID outcomes across the country by linking county-level estimates of measures of the preparedness index with county level estimates of COVID incidence and, in this case, COVID mortality.

This slide shows you results from those analyses showing you that places that score higher on the index generally have lower rates of death from COVID over this period of time. And in particular, looking at the sub domains measured in the index, it is that community planning and engagement domain that is

most strongly predicted with lower burden, lower levels of mortality, in this case, from COVID-19 after adjusting for demographic and social economic characteristics, including adjustments for the social vulnerability index that Dr. Mays mentioned as well.

And suggesting the kind of things we are measuring in this index are actually predictive of kind of how the COVID-19 mortality patterns have played out around the country. And again, community planning and engagement, the strongest element of this index in terms of predicting mortality from COVID-19.

Within that broader domain of community planning engagement, we have several different measures of community network strength, the strength of community networks that exist within communities. And we measure the strength of community networks using another national data source called the National Longitudinal Survey of Public Health Systems which basically measures how strongly public health organizations are able to work with a variety of other organizations in the community to plan for and respond to health issues in the community.

Working with medical care organizations, social service organizations, et cetera. We use some classic approaches to network analysis, network measurement and measuring the strength of these networks using density and centrality measures to do that.

And those network metrics that are included in this preparedness index are the strongest individual predictors of COVID mortality as we have done that analysis across the country. And making the case that these networks, the strength of multi-sector community networks, really matter when it comes to responding to COVID-19, as well as other types of large-scale hazards.

The last point I want to leave you with is to understand that these multi-sector community networks, which we have shown, are really important to planning for and responding to largescale health threats are also major users of race ethnicity data. And they are, in many ways, limited in what they can do because of the data limitations that we have been talking about in this panel.

These networks, we have done work with them. They rely heavily on small area estimates of race ethnicity from the American community surveys which we heard about earlier in the panel has high margins of error and other distortions in the data. They are also using administrative data sources, the kind that Dr. Bekemeier talked about earlier which are imperfect.

They are increasingly trying to reach out and use other data sources, the most emerging data source that these networks are using are social health information exchanges that are being built largely on top of existing health information exchanges. They are trying to integrate in data from social service providers, housing, food banks, et cetera.

Those are emerging data sources that we feel are promising again potential solutions to filling in some of the gaps in terms of race and ethnicity and vulnerable populations and ones that we should continue to look forward as a possible solution here.

I will end there. Thank you.

Denise Love: Thank you so much. And I am sorry to rush everyone because we could listen all day.

Next, we will go to Dr. Stoto from Georgetown.

Mike Stoto: Thank you, Denise, and everyone for inviting me. This has been a terrific panel so far. I have really enjoyed the presentations and hope I can contribute to it.

The work I am going to be speaking about, it draws in large part on the work I have did with a committee of the National Academy of Sciences, Engineering, and Medicine that was formed because of the controversy after Hurricane Maria trying to estimate how many people were impacted by it. And there were a lot of different estimates.

And as we were coming to the end of our work last March, COVID hit. And we realize that a lot of the ideas that we had been talking about in this committee applied to COVID as well. We ended up writing an appendix to the committee report that both illustrated what the committee was talking about, and then helped feed that back into the committee recommendations.

So I will speak about really the COVID aspect of this rather than the whole committee report. But the fences are there, and I have got some summaries of that as well for people who would like to see the more general version.

I have been focusing more like what Betty has been on is the kind of data that the public health officials need to make decisions. And it is the metrics. The metrics that we all see in the paper every day, that we want to believe are science based and objective and using them to make decisions.

They are published by a variety of organizations, government health agencies, universities, the media. They vary in what they present and how things are defined and what they deal with. There have been problems like uncounted deaths that already have been referred to by Andrew.

And it really seems that we are flying blind in a large degree against COVID. It is reassuring that actually the first element of President Biden's executive order about dealing with COVID has to do with building trust through helping better assistance. I think a lot of what I am saying will relate to that.

So this is the kind of thing that we all see every day. We probably get it in emails. This one is from the New York Times with the situation last week as I was preparing this presentation.

One of the things that pops up here, of course, is Oklahoma being blank. It probably was not true that there were no cases of COVID in Oklahoma, but some data system issue. The week before, Missouri was lit up like a red light. And that turns out to be because they decided last week to count all of the probable cases all at once at that time.

So there are all sorts of things like that, that happen with these data. It makes it hard to rally have trust in them.

These are important because states, counties, are using them to make decisions. California, in particular, I see Vickie shaking her head there because she knows that there is a very strict system that the governor has put in place with four tiers. And depending on where you stand, there are very different levels of restrictions.

I heard that San Francisco, people were cheering in the street when they went to yellow last week because of that. So it is very consequential what these metrics do. But can we trust them is the question?

So another way that we often see is the number of reported new cases daily. This is over the whole last year. Of course, it goes up and down. This was prepared last week. We know that in the last week, it actually has started to turn back up again.

But one of the questions that always comes up when you see data like this, is it because we were doing more testing or not enough testing? This shows the number of specimens tested. And it goes up. It was high in that winter peak. But is it high because they were many cases that needed to do contact tracing? Or were the cases reported because more testing was done? It is hard to sort those things out.

So sometimes people have tried to look at this by using the test positivity rate. Simply the ratio of the number of positive cases, the number of tests that were done. It is a very ad hoc solution. It depends on what kinds of tests were done, why they were done, and so on. I won't go into the detail there.

The key thing about this is that both the numerator and the denominator change over time in complex ways that really don't have to do with what is going on with the epidemiology of COVID. So it is really hard to tell what is really going on. This is an ad hoc and not particularly successful solution.

So why do we have these problems? I think it comes down to the fact that most of the data that we have, that we track, comes from case surveillance data. It is a standard public health technique that goes back more than a century.

Certain cases need to be reported to the health department who then use it to investigate or to do contact tracing. That is a complicated activity, particularly when you have thousands of cases like we are seeing these days. And it is hard to get it right. And it leads to a number of problems.

One that the epidemiologists talk about all the time is the iceberg effect. We know that only a fraction of the cases that occur, a fraction of the infections that occur, are actually reported. We like to think that, like a real iceberg, a constant fraction is above the waterline.

But that is not true because of changing of test availability, differences between tests in states in terms of what gets counted, different definitions and all sorts of other things, including many things we have spoken about here. This number just is neither constant -- the fraction that is reported is neither constant across states or over time. One of my graduate students has said, we really have a bobbing iceberg that goes up and down over time.

And of course, the same thing is true for reported deaths as well. And Andrews spoke about these issues very nicely.

So the other problem with it is that the people doing the contact tracing are focused on doing the contact tracing. And they don't always gather the information that we need if we want to study things like social economic differences.

So last month, when CDC wanted to publish some data on race and ethnicity disparities, they took data from only 16 states and big cities that had more than 70 percent completeness of race and ethnicity information in their case counts. Which is only 30 percent of the US population. And of course, even some of them, they don't have the data.

So this is a potentially very biased sample. It shows Hispanics and African Americans have higher rates. It is probably worse than that. But we just can't tell from data like this.

Another approach was also from MMWR last month. CDC because they don't have the data at the individual level, they said, let's look at the counties that tend to be one ethnicity or race to another and see which ones of those have high COVID rates. This is a very, in a way, convoluted way to get at what we want to know.

So what did we do? These are recommendations from the NAS report, as well as mine. So if we want to think about dealing with these -- working with these reported cases, we should standardize case definitions, as already has been done. But also standardize measurement processes and metric definitions and time periods for averaging and all of these other very practical things that are needed to make these data comparable. But there are limits to that.

In the initial Academies report, we spoke about basically building a system to do this better. And I don't have time to go through all the details here. I do want to talk a little bit about three different ways we spoke about in that report of improving the metrics.

Early on, we sort of have to go with what we have. But now, we are a year into that, and we think we ought to be doing better. And we want to be able to do better the next time we have an emergency of different sorts.

So we talked about three different methods that could be done. One is excess mortality methods. And Andrew has done a great job of that. So I am going to skip over five slides. I do want to say if we stop here, though, that there still needs to be work done in this area.

What we saw looking at the Hurricane Maria and the other hurricanes here is that basically, every time people come to it, they start from scratch and make up methods. But we think that there could be standardization here in terms of things like how do we do projected expected populations?   How do we take into account the fact that populations move?

During the peak in New York, some neighborhoods, 40 percent of people moved out of town. What impact does that have on excess mortality calculations. How do we take that into account? And many, many issues more like that, that if we thought about this systematically, we could use this as kind of a surveillance system going forward to track things, in addition to the kind of explanatory work that Andrew has spoken about so nicely.

Another one is the topic of syndromic surveillance which is an idea that goes back about 20 years. And the idea is that you don't actually have to wait for people to be diagnosed to be able to have something to be tracked.

And so an example of this is where we track people who either come into hospital emergency deaths or for outpatient visits and has been influenza-like illness, certain sets of conditions. But we could also look like COVID-like illness.

So on the next slide, this shows some data from the New York Metropolitan area last March. And we see the COVID cases in red that go up, a spike-up starting in March. The influenza cases in blue which were, of course, dropping off at the end of the flu season.

But the percentage of emergency department visits that were for influenza-like illness again was high in February and down in March and then back up again in March and April. Basically, picking up COVID

cases that were coming to emergency departments with symptoms that were like flu. And you could also look at this for -- this has been refined, and CDC is now tracking it for COVID like symptoms as well.

And then the other idea is that we can do surveys, based on representative sampling. Again, we don't have to count every case to make sure that they are valid. If we thought about developing a statistical system from scratch, we wouldn't think about training every doctor and every funeral director in the country to try to categorize things for us. We would develop a system where we focus on some cases and do it well.

One example of this is seroprevalence surveys. They could be population based as is done in some places. Or it could be based on things like people coming in to give blood for other purposes or in certain kinds of clinics like renal dialysis.

This shows data that was published. This is from last July of people who were having dialysis around the country in different places. And we see here in July, this was mainly picking up people with antibodies who had been exposed earlier in the year. So it was very much focused on the northeast and New Orleans area there.

But we also can use this because these are actually very large samples here to break it out and to look at disparities. So it turns out that the seroprevalence rates were substantially higher in ZIP codes that had high black and Hispanic populations and high level of poverty and high population density. And I suspect if you looked at the kind of depravation indices that Vickie was talking about, you would see the disparity there too if they had done that kind of analysis.

The Census Bureau set up a post survey early on based on existing sampling frame. And they were able to look at things like food insufficiency. The next slide shows how that varies by different racial and ethnic groups.

The key thing here is they were able to build on an existing population sampling frame. And they used questions that had been -- they weren't developed in advance, but they could be developed in advance. And what we are calling for in our Academies report is to have a system set up ready to go to do this kind of thing.

So the next slide is the final one. I think that managing the pandemic requires detailed objective data on a level and a rate of new affections. This better information starts with standardizing current case definitions, measuring processes, metric definitions. In other words, the things that we talk about as good research methods. In particular, having a constant reference population is really important and often missing in the current metric systems that we have.

But in addition to that, there are a number of research-based methods for doing estimations that can supplement and complement these counts that we look at so much. Excess mortality, syndromic surveillance, and surveys based on representative samples.

This last set of estimates is still largely experimental. And I said it tends to be one-off approaches. So research really is needed to developing the best methods, how to standardize these approaches. And also, we think there is an opportunity for the federal government to develop kind of a framework that can be put into place immediately after a hurricane or as soon as the pandemic starts. So we can really track these things in an ongoing way.

The last couple of slides give a little more detail about the report. But when you get the slides, you can look at those things. I will stop at that point.

Denise Love: Thank you, Dr. Stoto. We will go on to Ninez Ponce. And we are tight on time, so we will move right into her presentation.

Ninez Ponce: Thanks so much. Next slide please. If we don't make it to the end, here are the takeaways to the presentation. Data disaggregation equals data equity.

I am focusing on the Filipinx/a/o population within the Asian rubric and that invisibility of this particular population in the Asian aggregated category has implications, as well as implications of the invisibility of Native Hawaiian and Pacific Islander populations with this Asian Pacific Islander category, a residual category, are not reported that we see so often in government statistics.

There is also this under representation of Asians and NHPIs is one of this metrics that Dr. Mays had talked about earlier is the California Equity Metric called Healthy Places Index for vaccine allocation and other public policies, and then The Way Forward.

So data equity's definition is that data is collected, analyzed, and distributed such that the most marginalized communities have access and can use these data. So it is not about the average consumer having to do this, but the marginalized communities that have low literacy, limited English proficiency, do not have access to broadband.

It is a global imperative. It is not just in the United States. It is not just also about certain counties in the US. But this talk focuses on the US, Asian and Native Hawaiian Pacific Islander race and ethnicity data equity in COVID-19 which I have been tracking and following the past year.

So I don't think I have to remind people about the OMB Directive 15. But in 1997, there were revised standards with five minimum categories. I point out that Hispanic or Latino, not Hispanic or Latino, again, is a separate question. And then five racial ethnic groups. And that Native Hawaiian Pacific Islander are a separate group from the Asian category.

The OMB is a minimum of these categories. And that there is actually language in the federal register that encourages that we shouldn't limit collection to these groups. And further detail is encouraged, particularly if it is imperative to representative diversity of local populations.

So back on July 5th, 2020 was the reports on the inequity, the coronavirus. So this was The Fullest Look Yet back on July 5th. And the populations that were reported were whites, blacks, and Latinos for the entire population.

But back on July 2020, you can't say there was no data because at that time, there was data that showed that actually Native Hawaiian and Pacific Islanders, they are the dark orange line that is on the top, were already showing high case rates, case per 100,000.

And as you can see, as we continue to track this group through February, March of this year, that the case rates continued to escalate. And in fact, even along with the Hispanic population in California, continued to have a steeper surge than the other populations.

Another type of invisibility is that because public health systems do not report disaggregated Asian categories, that what we know about the pain that is suffered by some communities is only found in online tributes or GoFundMe or Facebook and Instagram. So this is a website of an online tribute for the Filipinx health worker population, not just in the United States, but actually globally.

As we can see, though, the United States is the epicenter of where the loss of health care workers are in the Filipinx/a/o population. And yet, it is not available in public-facing data. But it is available, like this insight is known from special surveys like the Guardian/Kaiser National data that lists the Philippines as the top of immigrant health worker deaths. From a National Nurses United Survey that 4 percent of the RN population are Filipinos.

But 31.5 percent of RN deaths are Filipinos. And then from my access to restricted data that took many months to get the 3 percent of the population in California of the population are Filipino health care workers and 31 percent of health care worker deaths, mirroring very much the National Nurses United Survey.

And in terms of case fatality rates, the severity, and I know technically, it is a disparate case ratio. But we are seeing that also the Asians do have the highest death case ratio among the major OMB groups at 1.18 percent. But if you disaggregate that, again this is my view of prescriptive data that Filipinos actually are the ones driving that at 6.4 percent case mortality rates.

Moving onto Native Hawaiian and Pacific Islander population, where are they disaggregated? And this is something from the NHPI policy lab that, with the NHPI community, I helped stand up during the COVID pandemic. And it shows here that it is only the orange slides where there is actually any data on the NHPI population. All the grays, they are not available. It is about 20 states.

When we think about this vaccine framework and the equitable premise and the criteria, so we know that from ACIP, that there should be aim to reduce existing disparities and not create new disparities. And from the NASEM report, that you want to ensure that there are special efforts in these high vulnerability areas, again now moving to this place-based metric.

And based on the Georgetown report, 18 states have used an equity metric for vaccine allocation. This was as of December. There may be more at this point.

Dr. Mays talked about some of these place-based vulnerability measures. I just point out, because I am going to show a map on the next slide that we do look at the social vulnerability index is the only index, at least in these four that I will be showing in the next slide, that explicitly have an input of minority status and language.

The Area Deprivation Index does not. The Health Places Index, which is what California uses, does not. The Pre-Existing Health Vulnerability Index, actually the medical vulnerability does, that Dr. Mays was part of. But the one that we have had access to also does not have minority status and language. And looking at multi-generational households and looking at limited English proficiency, for example.

So this is really hard to see, and I apologize. But the main point here is that this is a tool from the California Health Interview Surveys, a survey that I run. And it is called Ask CHIS Neighborhood Edition. And we have this at a tract level.

So we put all these indices in one place to see how it performs. I chose a tract in Carson, California. It is a high Native Hawaiian Pacific Islander population area in California. And you can see that there is only one of these vulnerability indices, the SVI, that actually tells policymakers that this is the thickest quartile, and that there should be some priority. The others do not. Others miss that.

I looked at under-representation then in using the HPI which is the Healthy Places Index which is in full force right now, not just in reopening, but in vaccine allocation. So looking at each county, I looked at the sickest quartile, the HPI, and looked at the communities of color that were under-represented.

The good news is the Latinx Hispanic population is not underrepresented by using the HPI. But the other communities of color are underrepresented, particularly Asian population and the NHPI population.

And it is underrepresenting however then juxtaposed with this overrepresentation in the cases index of NHPIs in particular than is really concerning for this community, for the NHPI community. When metrics overlook the need of a community because they are not geographically clustered because they are so small. And so these place-based metrics that so not include the social dimensions of pain for some communities then just really miss out the opportunity to address their needs.

I suggest that we enhance place-based equity algorithms. That they need to be augmented by other domains such as limited English proficiency, linguistic isolation, language spoken at home, per capita income, per capita versus median because a lot of immigrant households have large number of household members. Immigration status, frontline essential workers, proportion of multigenerational households. This is, again, also would get at the susceptibility and the spread of COVID.

And race and ethnicity is a public health emergency response. California's reasons for not putting race and ethnicity is political. But I believe it should be a public health emergency response to overwrite this. And then two considerations, do we model as a factor or do we layer this HPI with other criteria?

And I do like Glen Mays' community, engagement community systems criteria, because at the end of the day, then you are going to want to -- it is going to be mobilized through community effort. So the last thing is disaggregate, particularly for the Filipinx/a/o community. Thank you.

Denise Love: Thank you very much. We will go onto Dr. Cullen. But before that, we will extend the session here. I think our chair has given us another 30 minutes, if we need it.

If a panelist has to go, I know you were booked for a certain amount of time, feel free to do so. But we don't want to let this topic go just yet. And so we are going to extend our agenda time to, I think, 4 o'clock.

So Dr. Cullen?

Mark Cullen: I want to thank you, Denise, and the other organizers for inviting me to this fascinating panel and to congratulate those of you who have spoken before. I am learning a ton and that is what makes it worthwhile spending a day on Zoom.

I do, before talking about the solution that I am going to present, which I think only some of you are probably even vaguely aware of, I just want to point out the very different needs we have for granular and inaccurate data down to the level of the very accurate reporting of race and ethnicity.

So much of it obviously is for public transparency. Several earlier in the first panel talked about the critical nature for politics and everything else to be out of the way, so that the public can, on an almost daily basis, see exactly what is going on down to its social ramifications. That is one important use of this granular data.

A second very important use, obviously, is for the practice of both clinical medicine and for public health on a day-to-day basis. I am not going to be talking about those. I am going to be talking about a third need for extremely high quality granular data which is, for the purpose of doing the research which underlies many of the things that a year into the pandemic we now take for granted.

Because so much of this research, I think now over 200,000 papers in pre-print or print have appeared in the course of the year. And it is to this end that I think the consortium that I am going to describe to you represents at least a partial and interesting and, to some extent, unique solution.

The basis of what I am about to describe to you is, of course, not premised on anything the government does or anything any of our agencies, certainly not the interest either of physicians or public health workers in the country. But the remarkable reality of our health care system in the US which, because every dollar is counted and the system itself is driven by money, we track, not for any reason other than commercial, virtually every single activity that occurs within the health care sphere.

And it turns out that literally dozens and dozens and dozens of commercial enterprises ranging from small to massive are involved in this. So all of you know Epic and the large insurance companies and the like that are probably unaware of the many, many other companies that participate in the collection of data from where various forms of transactional care occurs, whether it is telemedicine or a visit or whatever, all the way out to the roll-up of those data that are mostly handled by very large organizations.

And it is in this vein that the very early onset of the pandemic, literally in March, many of these companies who interact with each other, both as competitors and as clients, got together before I was involved in it and formed a loose consortium, 30 odd companies, in which the goal was to pool all of their data, this is extremely granular data including the kinds of details that would make most public officials blush in terms of its privacy and security ramifications.

But to pool this data in a safe way and to develop a method by which that data linked together in appropriate ways could be made very rapidly available to the research community. And the principles were that the data had to be granular, that links had to be available, although never made available to researchers, but available, so that datasets could be effectively linked.

The quality of the data was important, so this had to be commercial grade data. It doesn't have, by any means, perfect reporting of race ethnicity. But it is far more complete than a wide variety of comparable publicly available datasets.

The companies had to commit to making the data refreshable on more or less a weekly basis, so that researchers would not be studying last year's outbreak, but last week's outbreak. And so the data are of that kind.

And then most importantly, developing a scientific review and administrative oversight using technology that can secure the privacy of these data to make them very rapidly available. So I am going to talk about the research database which is actually known as the COVID-19 Research Database.

You can see on this slide, the website. I would strongly encourage those of you who are not aware of it to go out to the website and to register, so you can actually see the full spade of both the datasets and the procedures. It is very open and transparent.

So this is just a sort of a structure of the way these so-called real world datasets are put together. Again, a three-part process in which a privacy security and compliance review, a scientific review, and then a proprietary review by the contributing data sources can occur within about a two-week period of time from a very short application procedure to having provision of the data in a secure and enclave environment.

So this is really just a schematic overview of how we manage this remarkable process. Just to put in perspective, there are now something like five or six billion records in there. And it covers about at least partial information on about 90 percent of the US population.

Unfortunately, there are pockets that are not available. So then, for example, people that are totally uninsured, undocumented and so forth, are less well-represented. But major swaths of the population at every level of their insurability are available.

So this is the structure by which the data are put together from a technical infrastructure point of view. I am not going to go into the technology that allows for the stripping and tokenization of the links, so that they are available to the data managers of the system, but never to any of the investigators who come on the platform.

So these are some of the key players. So among the 35 organizations that are now involved, almost all of them are for profit proprietary organizations. These six play unique roles. DATAVANT was the originator and the host and the original organizers of the database. HCCI, that many of you know, is a not for profit that has long experience in the management of health care data. And then three companies that are in the IT business in health.

So this is an example or an overview sketch of what is available in these data. I would not, for a moment, claim that these data are perfect in the sense they don't cover complete populations virtually anywhere.

And many of the datasets have varying degrees of quality for important fields, including the field of greatest discussion today, race ethnicity. Although not surprisingly, the private sector has taken this particular field, and many like it, quite seriously.

In particular, the datasets are linked to datasets containing social information in which these variables are very well recorded. They are probably not perfectly disaggregated per the previous speaker for which obviously there is a lot of improvement for future work.

This is what we have achieved in the first year just in terms of gross numbers. So about 2,200 people have registered for the dataset or 2,200 groups. 150 have made active research proposals.

And right now, there are 400 different researchers as of like literally today that will be working on the platform, each with their own unique area to manage their folders to access the extraordinary range of tools on the workbench. And to maintain data for review prior to any downloads that are done which we do internally review again for privacy and property considerations in that setting.

I will point out this is all completely free. The companies obviously are expending both personnel resources and money. AWS contributes cloud credits. Snowflake contributes credits. Virtually all of the companies provide some major personnel or other resources.

I am almost done. This is just to give you an idea of who is taking advantage of it. I just want to highly emphasize the first line. So we have had many applications from public agencies, partly because these data are richer than many of the available data sources in terms of granular details, especially in areas like health services and clinical data because we have tens of millions of EHRs linked to some of the claims data. But you can see that a wide variety of both domestic and foreign groups have joined this enterprise.

And this is just a quick overview of some of the more salient publications and other releases of data that have emerged from the study ranging from very granular studies of particular diseases to possible secondary uses of medications that were originally designed for other purposes, so repurposing. And then as I said, extraordinary range of health services related work.

So I will stop there. And hopefully in discussion, I can fill in some questions as long as you all are kind enough not to ask me too much about the IT technology because I take a lot of it at face value. Thank you.

Denise Love: Thank you.

Vickie Mays: I want to thank everyone for these stellar presentations. I hope that my colleagues are just brimming with questions which I have a feeling they are. So thank you for some additional time.

What we are going to do is ask you to use the raise hand function. And we are going to start with questions from the committee. And then move to others who are also in a position. And Rebecca is going to help us as well in terms of the questions that are popping up in other places, so that we can kind of try and stay focused.

Okay. Questions, comments?

Valerie Watzlaf: This was so fascinating and yet alarming, too. My question, I think, is for Beth. I think you said a few things that really resonated. And I think it was on slide 16, you were talking about imputation there. I just had a question as to how you decided to use that particular method that you chose?

Elizabeth Pathak: I think it is basically just the simplest method to use. And there is a very fundamental assumption behind it. The assumption is that the people for whom we don't have race ethnicity reported with their case status are essentially the same as the people that we do have race and ethnicity reported.

In epidemiology, we like to be conservative. So if we are going to make an error, we ant to err in the side of biasing towards the null. So you can make an argument, well, maybe race and ethnicity minority populations are more likely to not have their race recorded.

If we take the distribution of the known cases, and apply that distribution to the unknown cases, maybe we are underestimating the number of race and ethnic minority cases. But if we are doing that, we are making a conservative assumption, and we are biasing towards the null.

And I think the real problem is that we just don't know why we are not getting this data. Is it a problem at the level of the laboratories that are doing the testing? And I think it is not a scientific problem. It is a policy problem. It is a systems problem. But we need an investigation. We need to systemically investigate the patterns of missingness down to the local level.

Is it only certain counties, or is it only certain labs that are not reporting, and the other labs are doing a good job of reporting? So once we investigate, then we have the data that we need to try to change it, whether that is an incentive that needs to be paid or whether there is a consequence that can be given to say, hey, guys, you have got to do a better job on this.

But it is kind of depressing to me that, after all this time, and we have know that this was a problem for months and months and months, that it is still a problem, that it still hasn't been fixed.

So it needs to be more actively addressed. Not just talk about it, but someone needs to go out and actually do something and make somebody change what they are doing, so that we get these data.

Valerie Watzlaf: Just one quick follow up, and then I will be quiet. Did you also notice this and you were able to use that same type of imputation for other age groups as well? I know you said we should go to your website to find that. But do you --

Elizabeth Pathak: Our website is just the COVKID Project, so it is just focused on children and teens. But he problem is exactly the same for all ages. And the exact same method can be used.

I will say that the racial ethnic distribution of cases and deaths varies highly by age. And the younger you go, the more racial and ethnic disparities you see. Because in the elderly population, we have a lot of while elders in nursing homes and assisted are facilities who died from COVID-19.

So when you look at 65 and older and 75 and older, there is not as much of a racial disparity. But when you get down to young people, young adults, teenagers and children, and even infants, the disparities are very great. So the imputation has to be done by age and by state, whatever method you choose to use.

Valerie Watzlaf: Thank you very much.

Wu Xu: Thanks to all the presenters and organizers for having this informative and insightful panel. So I am most thinking related what committee's action come out of this hearing. So my question is, I am surprised and I need to know the census decision on which data elements to improve is a policy decision on the allocated fundings. So I want to make sure I heard right.

So far, there are focused on rate (indiscernible) and also the Hispanic population data quality improvement, right? So it is not at all the race and ethnicity data improvement. Is that correct?

David Van Riper: Not quite. I think the use cases that the Bureau has been privileging for the 2020 Decennial have been redistricting which is all race, all 63 OMB race groups, ethnicity, voting age, either above or below 18. And those are the three main ones.

And then the next set of policies is when we talk about race, it is going to be for all race groups. It is at the variable level. Seth was presenting on kind of Hispanics as a illustrative group for the ACS. But yes,

the Census Bureau is going to decide that we are going to allocate privacy loss budget to particular variables.

And because it is a fixed pot, if something gets as higher allocation, something else, by definition, will get a lower allocation. And that is a decision that is made by an internal Census Bureau committee who is taking input from stakeholders like the committee can provide a report on what they think is important. But the Bureau will have to balance out all the competing interests of all the groups.

Seth Spielman: Just to add, nothing has been set yet in terms of the 2020 Decennial and how they will allocate this privacy loss budget. It is a new idea. It does make explicit certain kinds of tradeoffs that really haven't been explicit in the past.

And I don't know that the process by which they will make the final decision is set. I believe there will be an open period for comment that they have recently announced.

Rebecca Hines: I just need to note that this committee reports to HHS and Census is in another department. So that is just if you wanted to make recommendations to HHS to do something with commerce, that would be fine. But just to make that clear.

Vickie Mays: I want to follow up on that. All of the NCHS datasets are based on census data. So it is almost like it is in the best interest, in terms of our data, to really think about what the implications are from census.

Seth, you started to actually go where I had a question. And that is have you or others started trying to model out the implications of the potential new privacy differentiation on specific racial ethnic minority populations?

Because I know for us in California, we are particularly concerned about the American Indians. I am sure that the smaller Asian and Pacific Islanders, and what that is going to mean for us in any data that has census as the basis for our estimates.

Seth Spielman: Yes. There was a workshop held by the Committee on National Statistics in kind of December. And those workshop proceedings have been published. Those were based on an initial release of this new approach. And both the videos and the report there are up. And there was a lot of looking pretty carefully at specific subpopulations and the implications.

Overall, it wasn't a great story. I think the Census Bureau took that to heart and worked on improvement. On Tuesday of this week, they just had a webinar in which they tried to sort of talk through some of those improvements at a high level. Those slides, I believe, will be publicly available.

They had a lot of different approaches, there were tradeoffs among the approaches. I think that I don't have a strong enough handle on the results yet to really summarize them quickly, other than that there are a lot of tradeoffs that they are trying to balance across it, I think.

David Van Riper: I will just quickly jump in and say that as of today, and in the Tuesday presentation, the Bureau said that they do not plan to release uncertainty estimates for the Decennial 2020 population counts, at least as of right now. That could change.

But once we know more information about the process, we will be able to roughly estimate uncertainty. But their methods are pretty complicated. The noise injection is pretty straightforward, but then they have to do a lot of post-processing. And that introduces a whole other set of issues that we can't easily model that uncertainty.

But I can tell you that that is going to be a hot topic of research over the next five years because it is so new. I'm sure Seth and others and I, will continue to work on that over the next little bit.

Vickie Mays: Can I ask you another question if I don't see a hand up? But after this, I want to make sure that the panelists get to ask each other questions.

This question is actually for Dr. Cullen. And it has to do -- I am aware of the data consortium that you are talking about. And it really gets into these new privacy kinds of concerns that I think people have been struggling with.

In terms of a review by the IRB, the IRB doesn't have jurisdiction over the mesh up. But this is under this all auspices of big data. And so the IRB doesn't really have the jurisdiction over this because it is deidentified. There is no PHII in here.

But, at the same time, populations, particularly racial ethnic minorities, have been concerned about for profits and all others utilizing this data. And for the small populations, what potentially could be publications that end up being somewhat -- that have a negative blow back to it.

And so part of this big data issue has been this notion of who is protecting and understanding racial ethnic minority issues? Who is protecting and understanding how data is used or sold? And so I am wondering if you could talk about this because this has been a discussion about the mix of that data. And many of us are kind of concerned that we not lose public trust. So those things that are going to make sure that that doesn't happen, maybe you could explain that because the public is very interested in this mashup.

Mark Cullen: Vickie, it is the $64 million question and I am probably not going to be able to completely satisfy you. Just by way of context, I spent five years at Stanford building -- not for COVID, this predated COVID or just came to the brink of COVID, trying to develop highly linkable extraordinarily rich, granular health data to make it available to academic researchers, both of our own and as widely as we could do it.

Under those auspices, we were beholden to the compliance and privacy offices of our own institution in the academic setting. As most of you in academics know, they are extraordinarily risk adverse. So many projects that could have happened, didn't happen, for that reasons.

The private companies have developed a technology that, at least from the purpose of the researchers using the data, their blind to any identifiers. And before they can download data, another review is done by the governance privacy and compliance group, that is intended to do much of what goes on at the FSRDC level, which is to look at cell sizes and to make sure that nothing is being reported that would result in a cell. The usual sort of cutoff number is any cell under 50 is not generally accepted.

But having said that, the standard is tethered to the legal risk tolerance of the contributing companies. And I guess the good news is that they are pretty risk averse in terms of that. So that we have lost a couple of projects at the compliance step for probably good reasons, although as the head of the

scientific steering committee, I am always pushing in the direction of trying to learn more, but we get pushback.

But it is a matter of public trust. I can tell you, having looked at the outputs so far, that I think we have done an incredibly good job in avoiding anything that would even come close to the worrisome line in terms of cell sizes and things that would raise those concerns.

But the flip side of that is we have not gotten into very small populations. So I am unaware, although several of our investigators are quite interested, for example, in Native Americans, and in some of the sub communities within the Latino community and also within the Asian US population, we have not had publications, for example, that look at Filipinos. Or that look at Pacific Islanders specifically, separating them out from other groups.

I mean, I think we will find ourselves in much the same boat that the whole field does. What we haven't done is try and distort the data through differential privacy or other forms of corruption of the data for the purpose of making it safer, but more useless. That is the tradeoff we haven't yet made. But this is a real conversation that we are not any of us going to be able to avoid.

Nick Coussoule: I have just one question. I am not sure who it would be directed to. But if we assume for the moment that all of the places where the data is gathered were perfectly willing to provide as much data as they could, is there enough information regarding what would need to be provided in education to make that happen?

I am trying to understand the differential between either a willingness kind of question, if you will, versus an ability and understanding to do that. And then some of the kind of technical semantic consistency questions that might arise.

Mike Stoto: Can I respond to that? From the point of view of kind of the case-based data that I was talking about at the beginning, that is generated by local public health people who are really busy in the middle of an outbreak. And getting them to generate data that they don't need for their immediate activity, doing it kind of in a research quality way is really hard.

I have been working with the public health people in Martha's Vineyard, a population of about 20,000, a substantial fraction of who are Brazilian. And it looks like there is the outbreak in the winter has been in the Brazilian population. But it is really hard to tell because they basically have stopped gathering the data that we need, the identification, as they got busy. So it is really hard to tell. That is why I think we have to find another way to look at it rather than those case-based data.

Elizabeth Pathak: I think there are a lot of technical and logistical problems at the state and local level. A lot of them deal -- the capacity to duplicate because you have multiple tests happening for the same individual. You have delays in reporting and there are a lot of issues with data.

And a lot of our surveillance is very date dependent, new cases, new deaths. But actually, it can take a variable amount of time for deaths to be reported, for tests to be reported. We know that there are situations where labs will rapidly report all their positive cases, but there are delays in reporting their negative cases.

So then this dramatically implicates the test positivity rate. So the test positivity rate can fluctuate on a totally artificial basis because of delays of labs and reporting negative cases to the state, not positive cases. So there is a whole range of really thorny, complex technical and logistical issues.

There are issues around definitions. Not all people who are COVID positive cases who die have COVID-19 as a cause of death. There is COVID-19 as cause of death, and then there is deaths among COVID cases. And those overlap to a great degree, but not 100 percent. And different states define a COVID death in different ways.

So I would say that it is everything. The more resources and standardization and kind of consensus on what we really mean about all these different measures are needed across all of the states.

Vickie Mays: Denise.

Denise Love: Yes, mine is down. It was answered.

Ninez Ponce: I think to get system change, you either have a mandate, an incentive, and you have guidelines on how to do it. So our issue in California is that we do know that there is collection of race and ethnicity at more granular letters than OMB. But it is not reported out by some local health departments and by the state health departments.

But yet, California has a code that you are supposed to collect more granular. And actually, in the education statistics world, they actually are told not to collect Asian, for example, because they have to go through the subcategories first. And that Filipinos are separate. So there are mandates and there are incentives with some of the COVID -- going to counties. If some of that could have been tied to either better, more granular data collection, and I am talking about my world of data disaggregation here. Or again, like partnering with representatives of community organizations that would really say you really need to do this. You really need to look at this population that is overlooked.

But standardization, so having guidelines on how to do it, and there are guidelines. I mean, I worked on a guideline for NAM in 2007 on guidelines for standardized race ethnicity reporting, collection and reporting back in 2007. And it talks about the rollup as well.

And surprisingly, it probably isn't statute that a lot of states are supposed to collect more than OMB guidelines, but there is no appropriations money where there is no incentive.

Margaret Skurka: Quick question. I heard discussion about cause of death. But on our death certificates in this country, the bottom line is underlying cause of death. And I worry about the data that has been happening over this past pandemic and how consistent that is statewide or nationwide in terms of underlying cause of death.

Because it is like saying, heart failure. Well, how didn't die from heart failure? But it was due to, due to, due to. So I don't know if any of the panelists want to comment on underlying cause of death for COVID.

Andrew Stokes: I will defer to some of the others on the panel because I see several people raising their hands. So let me defer to the experts here.

Elizabeth Pathak: I have done a lot of work with mortality data over the years. And Andrew, I really appreciated your presentation and mentioning the issue of home deaths. And of course, part of that also is outpatient, what is coded on the death certificate as outpatient emergency department deaths.

And this is a large portion of COVID-19 deaths. And as you know, when a patient comes into the emergency department, they can only stay there for 24 hours. And then they have to be admitted or they have to be sent home. So if someone dies in the emergency department, it means that before they crossed the threshold of the hospital, they were already actively dying.

So this is a failure of medical care and a failure of rescue when we have people dying in the ED. The ED is not a place where people should be dying.

And classifying cause of death is very difficult when we have limited medical information. I know Andrew used the term, garbage code. But I just want to push back on that a little bit. I know, Dr. Murray at UW has popularized this phrasing of garbage codes.

But I think it is problematic because the codes that they are referring to, which are symptoms, signs, and ill-defined conditions, is the family of codes when there is inadequate clinical information. So it is not really garbage. It is that whoever has to certify that cause of death does not have enough information to know why the person died.

And when people with lack of access to medical care, maybe they don't have a regular physician, when they die at home, or they die in the workplace, or maybe they die in the ED before the family can arrive and any information being given, the person who is writing that death certificate really may not have enough information to say why they died.

They may have some suspicion. And then people will vary in terms of whether they will just say, oh, well, we don't really know, so let's call it heart disease. And that is a very common thing that happens.

Or they will use these symptoms, signs, and ill-defined codes. But I will say, in direct answer to your question, when researchers have access to death certificate files, where they have all the -- so not a tabulation. And all we have right now from NCHS are tabulations. But if we had access to actual death certificate files, you will have both the underlying cause of death and contributing causes of death available for analysis.

So that information does not have to get lost if you have adequate data to look at. You can look at both and see whether some people are more likely to get that contributing cause versus the underlying cause for COVID-19.

Vickie Mays: Before I turn it back over to the chair, I just want to see, Rebecca, do we have any comments and questions that we need to take from our Q&A?

Rebecca Hines: Vickie, I think I would like to read into the record one in particular, from Ruth Wangerin (ph.). She wrote, I rarely hear anyone explain the limitations of comparing statistics about COVID that are derived from testing for infection.

Rarely are differences in testing coverage from one neighborhood or race ethnicity or age to another reported. Yet cases and present positives in these different groups are compared as if they are comparable.

Unless the population for a ZIP code, for example, are being regularly and randomly tested and across all age groups, how can we interpret statistics like number of cases or cases per 100,000 or percent positivity in the population in that ZIP code?

Clearly, the US never got the ideal surveillance system going. Yet even professionals talk about the statistics as if they are real. In some populations, people are rarely tested even if sick. In other populations, people are frequently tested just to make sure they are not infected. These differences in testing are related to race, ethnicity, and especially socioeconomic status.

And then she continued, the test positivity rate is being interpreted as a measure of prevalence in the community. But it is not comparable from one community to another because it depends on rate and thoroughness of testing coverage.

Vickie Mays: Any comments from our presenters?

Ninez Ponce: I absolutely agree with that. Thank you for that. It is also some populations are not getting tested because they fear for their lives for being blamed for the Wuhan virus. We are hearing that there is undertesting in the Asian community.

Melissa Goldstein: Rebecca, with the remainder of the comments, there are many more actually, be read into the record as well?

Rebecca Hines: This is an anomaly of Zoom. Normally, if we were in public, this wouldn't be possible. So I think there is one I will read when we get to public comment. But other than that, I think that is the one that is relevant at this time.

Melissa Goldstein: I think if we are going to read some into the record, they all need to be read into the record at some point during the meeting.

Rebecca Hines: Well, we can append them to the meeting summary, Melissa, in terms of being legitimate.

Vickie Mays: I was just going to say, let's just see if we can accept them, and that they will then be in some format for us.

Rebecca Hines: The official meeting summary that gets signed off on and posted on the website.

Vickie Mays: Great. So that it will be available.

Denise Love: Tammy has her hand up. Do you have one more, Tammy?

Tammy Banks: We spent many years trying, when we were looking at health care standards, to really understand that clinical use cases and administrative use cases use the same types of data. And when I hear all these presentations which were amazing and scary at the same time, it seems like it is the same data.

And is there any efforts or any consideration of figuring out what is the base source for certain types of data elements, so we can increase the accuracy? I mean, I know race and ethnicity is collected in a lot of different formats. We heard real estate here, we heard there are a lot of different industries.

How do we harness this data to increase the accuracy for research, public health, health care, real estate, all the different types of sectors because this is basic data that is very difficult and expensive to collect. But again, this is just a naïve comment. I would love just some guidance on that.

Vickie Mays: Ninez, do you want to answer?

Ninez Ponce: First, I would start with I don't think it is expensive. I bring education statistics again. I think it is an imperative that we continue to collect race and ethnicity.

But how do we do this? I mean, do we do it as granular as possible in terms of collecting? And then, of course, because honoring disclosure risk, honoring that and then we present data -- as possible for as many groups that are reliable and that ensures that there is no disclosure risk.

I am not sure if I am really answering your question about how like do we save the public health system. But I am sure there are other panelists here that could comment.

Tammy Banks: I see this being more than public health and more than research. It is actually all industries who use this type of data. There are a lot of dollars. And if we are putting a lot of dollars in all the different sectors to collect this data, it is an annoyance for each of us who have to put it in several times, right?

How do we harmonize across these industries that need this data, to get our increased accuracy in the research and public health and health care where it is needed? Not that the other sectors aren't important. I didn't mean that.

Denise Love: I am going to just say that I don't know about across sectors. I can't get my brain around that. But I have always had a dream that in health care, we could do a better job of centralizing the collection. And then having that carry through the health care.

And one of the examples is enrolment. That is a low stress time. It is when somebody can self-report their data. And just like at the birth record with the mother, those are potentially useful sources that could, through linkage and following the health care claims, flow through. But we aren't doing that right now. So that is just my two cents. I am frustrated that we can't do it at the point of enrollment at least, and start carrying it through.

Tammy Banks: That is interesting because if it would be in enrollment, there is less of an incentive not to be accurate because there is not a disincentive for answering in a certain way, right? Hopefully.

Denise Love: You can argue, there is always a disincentive and always a concern. And I think that is another thing is culturally, we are overcoming. But I won't get into some of those discussions I have had over the years.

Wu Xu: As a follow up on the discussion, and we heard on the incentive several times from different panels and the committee members. So for incentive, through the health care providers' data collection, I wonder, we have Medicare and Medicaid to tell us how their race and ethnicity data in their system.

I think that they really have a power to give incentives to the payers and providers to get better data. So if payer provider working on the Medicaid Medicare enrollees, overall, their system information will become more complete and accurate. So that is a comment about the incentive part.

Tammy Banks: I hate to say patient ID, right?

Denise Love: We don't have time for that. Glen, you are up.

Glen Mays: Just a brief comment. I think there is also information lost in the kind of chain of custody as health information gets exchanged, when you think about provider who orders the COVID test may have race ethnicity well documented in electronic medical records. That order goes off with the sample to the lab. Could be the public health lab, maybe a private lab.

That is an opportunity for that information to get lost. And then when the results get reported back by the lab to the public health agency, that is one place where race and ethnicity data just get lost. It gets separated from the other pieces of information.

Vickie Mays: Good point. I see no other hands up. I think at this point, what I would like to do is thank our participants. Our cup floweth over in the sense of the great information that you gave us for us to really think about, and the specifics that you provided of kind of what some of the gaps are, as well as what some of the solutions are.

So on behalf of the committee, Denise and I thank you for the seriousness by which you took your duties of providing us with some direction and not just identifying problems. I will then turn it back over to Nick.

Nick Coussoule: Thank you, Vickie. Again, I want to thank all the panelists and participants. Very informative. Someone else used the term earlier. It was really interesting and really scary at the same time. But thank you again.

Now, I think we have got a few things left to do today. But we have all been sitting here for a while. So I think we need to take a break. And we would reconvene at 4 o'clock to wrap up a couple of remaining topics we have, and then provide some time obviously for public comment at 4:45. But we will now take a break officially until 4 o'clock. I appreciate everybody being back on time, so we can wrap up our day. Thanks again.

(Break)

Nick Coussoule: Welcome back. We will get started again. We are going to do an audible and switch the quarter of our last two committee sections. We are going to cover the workplan first and then whatever time we have left until public comment we will come back to our Report to Congress, hopefully focused on big and substantive things and try to leave some of the more detailed editing for later on. If we could bring the workplan up please, Rebecca or Kim, whoever is going to pull that up for us.

Rebecca Hines: I will pull it up.

2020 NCVHS Workplan

Nick Coussoule: I will walk through the first part of the high level and then I will comment on the additional anticipated plans and activities by each of the Subcommittee chairs and then like each of them to also just briefly walk through the things that are on our parking lot of potential topics just to recognize what we have considered and I believe, are important to at the Subcommittee level, but may not have put in our actual product or workplan yet.

If we look up our Full Committee meetings schedule, we have one scheduled for September, which is our last fiscal year of '21 Full Committee meeting. I am not quite sure at this point in time whether that will be virtual or in person. We can all hope, but we will just have to see how things play out a little bit. We will obviously let everybody know as far in advance if we can regardless of that status so that we can all be prepared for that.

We also have time allocated for a couple of one-day subcommittee convenings. We have already talked. I think both of the Subcommittees talked about the need for that this year. I do not think we will have any problems filling it. We are just going to need to figure out how to sequence it within the calendar as well as making sure we have enough time for all those involved to adequately plan.

As far as the planned and anticipated activities I'll cover the first one, the Report to Congress, obviously, we know what we are doing today trying to get through that ideally in the second quarter. We will try to finalize it hopefully. We may need to hold some kind of a special meeting to approve. We will work on the details of that but try to get the editing continued done through the Subcommittees as well as through the coordination of a writer and then try to hopefully get to a final product pretty soon and then we will just figure out the mechanics of how we need to formally approve that.

Moving down then I guess what I would ask is that the Standards Subcommittee, Rich and Denise, to cover the new few items specific to Standards as well as to skip forward just a little bit and cover the potential topics. Actually, we will go through the actual topics first and then we will go to the potential ones.

Rich Landen: I will take a stab at that, Denise, if that is okay with you. Line 2, standardization of information for burden reduction, aka, Convergence 2.0. We spent a lot of time on that yesterday. We have consensus around the project proposal. The next concrete step is to start planning the industry listening session. The Subcommittee will be doing the planning for that.

Over the next month – I guess today is the first day of the second quarter, the actual listening session itself we have here penciled in for the third quarter. We will try and do that as early in the third quarter as the planning will allow.

The other sub steps in the first quarter reviewing this and the project proposals, and we have as a third bullet indicates we have integrated a lot of new concepts into the project scope that were not there with the previous projects scope for 18 months ago.

Far right corner, Standards is the primary owner. There are some overlaps of course with the Privacy Subcommittee. We will be touching base as necessary, as we get into the areas where there is overlap between the two subcommittees.

I think that's it on the project scope.

Rebecca Hines: Jamie has his hand up.

Rich Landen: Jamie.

Jamie Ferguson: Yes. Thanks. When we talked about this yesterday, we talked about an environment scan, which I think is the listening session. Then we also talked about getting stakeholder input on priorities of the sequencing, which I think is different from the actual environment scan.

I just wanted to open up the discussion a little bit about what are you thinking in terms of the listening session. Would it cover both of those different topics or might we potentially consider an RFI or some other mechanism, in addition to a listening session for the environment scan versus the priorities and sequencing?

Rich Landen: Excellent questions. I think those – you are going to be part of the planning process rather than scoping it out here on our timeline. I probably understated what I was really anticipating – all the planning for the project. We will now, as the Subcommittee, commence now that we have the project approval. And for the information for those who are not on the Subcommittee, the Subcommittee meets every other week. We will use our biweekly meetings to do the planning and that planning will of course, be based on the conversation that we had yesterday.

Part of it is if I recollect some of the key takeaways for planning purposes from yesterday, is we have identified some specific groups and organizations whom we need to connect with earlier rather than later to inform our planning, and then after we do those connections then we can pin down the exact scope of the listening session and who is to be invited and what we will be focusing on there, which were your specific questions, Jamie.

Anything else?

Nick Coussoule: Rich, can I interrupt for just a second? I know we have some people with some time constraints on the PCS Subcommittee. If you don't mind if I could skip forward to them and then we will come back.

Rich Landen: Of course.

Nick Coussoule: Jacki, if I can let you start on this.

Jacki Monson: Yes, I can start on it. Can we scroll down? Line 6. Two things that are on our workplan. One is what we spent a lot of time this morning discussing which is improving security and health care projects as well as that timeline.

The other thing that is on our list of things to do once we get through the congressional report is to flash back to September 14th of 2020 when we had the virtual hearing related to COVID and some data collection and the desire for us to create a set of recommendations or perhaps a letter to the Secretary with respect to that. Those are the two things that are on our workplan today.

Maya Bernstein: I know that Rich mentioned that there was a project that might have some cross pollination with the Privacy group. If you wanted to discuss that at this time too, maybe if you wanted to add that as something where there is cross pollination and the Privacy group might be also involved, that might be useful.

Rich Landen: I am not sure, Jacki, that we have a knowledge about at what specific point we will be getting into those issues. It is mainly around the privacy of the new or the approaches we will be considering for information flows.

Jacki Monson: Maybe, Rebecca and Nick, we can add a little placeholder because it might make sense. Will you go down that road to lend you a couple of PCS members to participate to help with that evaluation and standards.

Nick Coussoule: I think that makes sense. We obviously want to open up to any committee members who want to be able to participate across the initiative. I think that will be helpful, Jacki.

Lorraine Doo: I can tell you one of the areas that HHS is looking at – I do not know, I have to look at ONC, but for some of the CMS policy decisions and mobile applications, as we were talking earlier, and release of data is the opt-in/opt-out discussion, and patient privacy and some of the implications of those decisions and secondary uses. That is where some information or guidance or input from NCVHS might be interesting.

Nick Coussoule: Jacki, do you want to comment at all on the parking lot items for PCS at this point?

Jacki Monson: I can just comment generally. We still believe – we looked at the parking lot at our last couple of meetings and we still believe the things on the parking lot are important. But at this time, I think we have enough on our plate addressing both the security piece and going back to the September hearing, that they will stay in the parking lot for now, but we want them there purposely because they are important. I am sure we will revisit them when we have moved on from our current workplan.

Maya Bernstein: Jacki, this is Maya. I am just wondering about the timing of days. I am wondering if we could maybe somehow move up the work on the September hearing during the time that people are editing or whatever we are doing with the Report to Congress. I am a little worried that it will languish too long. People will spend a lot of time re-reviewing what we did in September because it is too far away. But also, I hope that by the end of this meeting or the next Executive Subcommittee meeting, we will be closer to what we are going to – or the next Privacy meeting – will be closer to what we want to do with that report and can be editing the parts that we have already talked about so that we can maybe do them parallel tracks.

I do not know if that is feasible or not. We have to somehow first with a couple of extra staff members now. Do you have any thoughts about that?

Jacki Monson: I think it depends on the progress that we make on the congressional report. What we heard from the committee members at the last meeting is it is hard to focus on more than a couple of things at a time and because we are going to be moving towards preparing for our hearing on security as well as a congressional report. I think if we are in good shape with respect to that soon in the next couple of weeks, then yes, we should be able to pivot towards that and leverage our interns to assist us in moving down that path.

Rebecca Hines: And Jacki, is your thought on the follow up to identify some recommendations and define a letter? Is that really what the project entails?

Jacki Monson: Yes, that is what we heard from Sharon would be helpful and I think the Subcommittee has agreed on that. We are in this transition where we only have a couple of members participating. We

have lots of new members and now we have interns. We have some work to do, I think, to refresh the hearing and what we learned from there and then proceed forward. Yes, that is the plan right now.

Rebecca Hines: The timing is great because there is April, May, June – there is five months, which is a reasonable amount of time to get recommendations and a letter drafted for September. So far, for best-laid plans, but right now, your timeline looks reasonable.

Nick Coussoule: Thank you, Jacki. Go back to Rich again.

Rich Landen. Just one more comment on Line 2 on the listening session. Every member of the committee will be invited to attend that. It will not just be for the Subcommittee members, but attendance will be at your discretion. We, of course, do not know at this point whether it will be in-person or virtual. We are presuming virtual, but that is no longer certain thank goodness.

Rebecca Hines: I would just like to invite both Subcommittees that the sooner you can start doing the calendar dance on both of these, getting them scheduled, the more likely obviously people will be available, members outside the Subcommittee will be available. Just another task but whatever assistance you need from Marietta and myself on that, just let us know.

Rich Landen: Thank you, Rebecca.

Line 3, Collaboration with ONC and HITAC. Again, going back, referencing yesterday's discussion with Micky Tripathi, it looks probable that there will be an ICAD 2. The calendar is open. There is nothing concrete or even preliminary to schedule on that, but we are holding the time open. And the Subcommittee understands that at some point in time, there may be specific activities and we will add them to this line item as they become known.

Line 4, the Predictability Roadmap. That activity is really flowing into and being subsumed now by the project proposal for Convergence 2.0. The concepts of the Predictability Roadmap will be pursued under the project on Line 2.

Line 5, this is a standing obligation of the NCVHS and specifically the Standards Subcommittee. If we get requests either from the designated standards maintenance organization, which would include requests from the SDOs, X12, NCPDP, or Health Level 7, or if we get a request from the Operating Rule Authoring Entity, CAQH CORE, that we will respond to in the time parameters that will be dictated by the nature of the request.

At this point in time, we do not have any specific information. We do not have any information about any impending request, but again, this is a standing item. We have to make allowance for it and react to it timely if it comes in.

Line 7, Evaluate ICD-11. Again, we spent considerable time on this yesterday. We sent a letter a little over a year ago to the Secretary. Right after we sent that, the pandemic hit, and all bets were off. Since then, of course, there is a new administration. We now have a new Secretary. We think that in the time between now and our September meeting, the Subcommittee will be able to take a look at the conversations and the presentation by Dr. Kin-Wah Fung yesterday, and take a look at our letter and dust it off and maybe re-issue it or send an update to the Secretary.

We would bring that draft through the Executive Subcommittee for review and action at the September Full Committee. That is assuming we do not get a response to our original letter by the new Secretary or someone in the department prior to that. If we do, obviously, we will pick up on the response to our original letter. But either way, this is the line item. This is part of the work that is planned for the rest of this year.

Denise Love: I have a question on that and I agree so far with everything, Rich. On the ICD on the NIH evaluation, is that a one off because it is such a small subset of the codes that were studied? Do they plan to keep studying or is that just a one-time study? Does anyone know?

Rebecca Hines: I would suggest that we have a follow-up conversation.

Rich Landen: Yes, we need follow-up conversations. I know we still have access to Donna Pickett. We can work through her just kind of a general let us see what was in our letter for the Subcommittee and what has changed since then and what are the hanging chads that we – or the new chads that we need to focus on.

Denise Love: It is just my curious mind. Thank you.

Rich Landen: Good. I think we have a couple of items in our parking lot.

Jamie Ferguson: Sorry. Back on ICD-11. My sense from the discussion yesterday was that if we wanted to have a follow up on the specific options that were recommended for modification or extensions then I think we would need to make a new request.

Rich Landen: I agree with you. That was new. That had not been under consideration and that is the – what we got yesterday in that sense is kind of a partial answer to the list of research questions that we had recommended.

Rebecca Hines: It seems to me that that is part of the letter to the Secretary that we can observe. There has been some research done and we engaged with it. What are the plans for HHS – the remainder of this that needs to be surfaced so that a decision on how to deal with morbidity especially moves forward?

Tammy Banks: Just a quick question. The last time we talked about ICD-11, it seemed like there was an open window for consideration of change. Is that window still open? Is there a higher priority for us to move this faster so we do not have to have a clinical modification, or does it appear a clinical modification is inevitable?

Rich Landen: I do not know the answer to that question. I think that is one we will have to get Donna Pickett involved with.

Margaret Skurka: If we want to have a little taskforce or a little subcommittee, certainly Valerie and I from the HIM profession would be happy to be on that to do some work there. I am not just suggesting her without – we have chatted about this privately.

Rebecca Hines: For the record, there was basically a subgroup of members ad hoc from Standards and Privacy that drafted the previous letter to the Secretary and the two attachments because it does not require a whole subcommittee. That can then be brought to the two co-chairs of the Standards

Subcommittee for the Full Subcommittee to review. But I encourage a small group of people with the expertise and time to do that work, Margaret. Thank you for organizing that.

Tammy Banks: I am happy to help, Margaret, if it is helpful.

Wu Xu: I wanted to volunteer to do that.

Jamie Ferguson: Me, too.

Rebecca Hines: We have lots of volunteers to work on this.

Participant: And Jim too.

Participant: Good. We want the physician perspective.

Rebecca Hines: Fantastic.

Rich Landen: It is almost a committee of the whole.

Nick Coussoule: Deb, you had your hand up.

Debra Strickland: I just wanted to volunteer for the group as well.

Nick Coussoule: Volunteer is always a whole lot better than voluntold. That is great.

Tammy, you still have your hand up. Margaret, is that just hanging there?

Rich, you wanted to go through the parking lot topics briefly.

Rebecca Hines: Before we do that, the follow up from today sort of falls into this general bucket. It is a different project, but it is the same kind of topic. I did not know whether we wanted to talk about next steps. I think there is a question in the chat about that.

Vickie Mays: I will comment if not. I was not sure exactly what was in the chat. I am going back and forth between the two. One of the things to say about the panel today in terms of following up on that is prior to that, I will tell you that there was a lot of HHS interest and different branches of HHS that were like they sense that. I know that there were note takers from the Surgeon General's office. There were note takers from OMB and several other offices. Whatever it is that we were covering, it is clear that there is the desire to get information and learn more about it.

The question is should we pull something together from what we heard and if so, how we might be able to do that.

Denise Love: Vickie, I was thinking the same thing and I think we should.

Sharon Arnold: I would agree. It was a really fabulous panel. I would really like to think about whether there is a summary and next steps. Certainly, there were lots of areas that made sense for the committee to focus on in terms of data standards, recommended standards at the state and local level, what to do, how to adjust for missing data, kind of methodologic issues. There were just a tremendous number of things that could be followed up on.

Nick Coussoule: I think two things I would add. One is I think we would treat this similar to if we did a day-long hearing in some cases that we might want to summarize and report out what we had. This is probably a big enough topic with enough interest that would warrant doing that, again, without volunteering people to do more work. But I think that is very valid.

And the other one is we talked about items that are on our work list and potential to-do list, but we always want feedback from our federal partners, obviously, from the Secretary on down to the people that we work with and try to support and provide some advice to the priorities and things that are challenging that we believe and they believe would be in the committee's purview to be effective then and try to figure out how we make sure those priorities also get worked in the workplan. We will always keep our ears open for that.

Vickie Mays: There were pieces of what we heard today, that to some extent I think could get put into Standards or PCS. I think that Sharon may have particular things that she wants to see, but there is a way to go through and look at it and also say this is the expertise of subcommittees we already have standing since we do not have Population Health standing right now.

It reminds me in the past where we have done some things where it is the committee as a whole that takes the issue on. But given that we are just getting some of the people appointed and up and running, Sharon may also have a sense of priorities of things to be pieced out and other things that might go into the Subcommittee buckets.

Nick Coussoule: I think that makes sense, Vickie. We have also in a few cases at least in the six years that I have been here, we have charged certain individuals with leading an effort that may not be one of the Subcommittee chairs, but we roll it up that way from a coordination communications standpoint. I do not want us to get caught into what I would call some arbitrary structural question more so making sure, one, we have capacity that our members and staff and support would be able to undertake something. At the same time, it does not mean we cannot start something else if all those things line up appropriately without over burdening obviously the members to not be successful with what we are undertaking.

Denise Love: What is a timeline for a transcript because that would be a starting place?

Rebecca Hines: They have ten days to get those to us. I would say by early the week of the 12th we will have the transcript.

Denise Love: Vickie, maybe I or you, could take a stab at distilling that. I am just trying to think of next steps that are logical.

Vickie Mays: I think next steps might be – because I think there was going to be a video because I think there is even more people who want to watch it who could not attend. If we get that up and then look at the comments and then start – I think we could start to go through it. But it also might be useful for Sharon to give us some sense of particular issues, unless she wants us to pull the issues out, but otherwise, I think that – I know different parts of HHS that are wanting to work on this or are working on it. But it may be that Sharon has a sense of a priority to us to help in a particular way.

Sharon Arnold: Let me talk to a few folks. I have some ideas. But let me talk to some folks that were also listening in to the conversation and get back to you with a sense of what the questions are from HHS.

But I think you should feel free to think about what you think are appropriate next steps as well. We can talk about that.

Vickie Mays: It sounds good.

Denise Chrysler: That was a fantastic panel. It was so good. I was torn between trying to take notes and more carefully listening. And I am very relieved that it will be in video as well as just a transcript.

But for my notes, and just thinking of what do people need or what do federal agencies need, I was trying to note innovative solutions. I was trying to note all the different factors that were trade offs in the margin of error. This is an area I do not know. But being a lawyer who is always trying to figure out the de-identification piece and the granular and the usefulness versus reducing risk of de-identification, there were so many options suggested that people like me in the field just do not get exposed to and how one provides an easy resource of the innovative things that are happening and the considerations and thinking of a Step 1/Step 2 and trying to make your decisions like what is your aim and do you really need data that is that granular and what kind of identification avoidance techniques would really work. That is how I reacted and I do not know if that would be something that would be useful to others.

Tammy Banks: I was just going to go the same way as just a fact sheet with a summary. Here are the issues, inconsistencies, recommendations so that we could use that if you did choose to do a letter of recommendations to whomever we would be able to make recommendations to.

Vickie Mays: There may be two different things to things about here is what to recommend if there is going to be recommendations. But given that we are starting to have videos and stuff is a way to package what is relevant to different groups. Because one of the things about the hearing in terms of how it was pulled together, is it wasn't down a committee line. There are things here that are standards. That are things here that are methods. There are things here – so it kind of doesn't fall for one group.

One of the things at least in COVID that we started doing is often taking our webinars and repackaging them and pushing them out the door to specific lines of people. It might even be that, that if you are doing de-identification of data in COVID, here is where you go. Here is for the kids' people because everybody is not wanting to sit through three hours of – I have a four-hour webinar. People did not want to sit four hours so we pulled some things out and made little things for people. That is another way to go if we want.

Denise Love: And then identify a few unresolved issues, which –

Vickie Mays: A lot of those.

Nick Coussoule: One or two of those. Tammy.

Tammy Banks: Is this another small committee?

Vickie Mays: Is it the committee of the whole?

Tammy Banks: A small committee of the committee of the whole. I am more than happy to pull fact sheets together, but I am not an expert. I can only pull together what resonates with myself.

Vickie Mays: But that is important.

Tammy Banks: In terms of the people that you know will then respond. Just like what Denise is saying, the attorneys want this. Somebody else wants that. We are learning that to the extent that we can push out what the people need and not ask them to do everything then we are willing to then read it and listen and use it. Here are three hours for you to look at.

Tammy Banks: Is this something that we can designated a few people and then get educated by Rebecca or whomever, on what we can do in regard to dissemination of this information and preparation obviously to understand what was conveyed so that we can get to those recommendations at some point, which would be the end goal from a committee perspective?

Nick Coussoule: I do want Rebecca to comment on the FACA stuff obviously. We have published reports and findings – necessarily directed letters at the Secretary for recommendations, but tend to be what I will call in progress work product, if you want to call it that and those kinds of things.

Rebecca Hines: No, I was just going to say that Sharon made a comment recently about if there is a receptor site, if there is some place within the department that is looking for specific advice, I think we would be well to hear what that is in the next few weeks. That would help guide what the next steps are specifically.

Sharon Arnold: With respect to race and ethnicity data, there is a huge effort within both HHS and the federal government to report everything by race and ethnicity. And what we heard today is that data is not perfect. You need to consider what the implications are. There was a tremendous amount of information that we could bring to bear on all kinds of efforts to do that where people may not be expert epidemiologists or statisticians and really know how to make those adjustments so what do they need to consider. I think that would be universally welcomed across HHS.

For example. If we think about what kinds of efforts, we need to think about going forward in terms of if we are going to collect additional data, what should the priorities be? We cannot do everything. What are the most important data elements we should consider going forward? Those kinds of questions I think are pretty universal within HHS. But I can certainly pull my colleagues to see what perspectives they had on this and if there are additional questions.

Vickie Mays: Also, I thought the original kind of question or request, is what we could shed light on that is relevant for states or local entities as well. Some of this is not just at the federal level. There was purposefully a lot of state issues and in particular, we rarely give advice to the rural areas. You can see rural areas are – they have a capacity and a workforce issue. I think there are some state issues here as well.

Denise Love: Yes. That is one of the unresolved issues.

Nick Coussoule: Deb, I see you have your hand up.

Debra Strickland: I was just volunteering to help distill some of the presentations down. If you are going to create a little group or a big group, I would be happy to help in that regard.

Vickie Mays: Thank you.

Denise Chrysler: I would also love to be on a group of this sort.

Tammy Banks: Vickie is nominated as the lead with Denise.

Vickie Mays: -- the work I am doing. I have no problems – you can see, I gave you what I thought were really good presenters.

Denise Chrysler: I had volunteered to join the group. I wanted to add when we are thinking about focus and recommendations with regard to federal agencies or HHS, they get a lot of their data from state and local governments so how it is collected by state and local government and the quality of that data just means so much within the Federal Government.

Denise Love: Right. And to follow that, for AHRQ, I have done some assessments on some of the administrative databases. Some of that is already gathered and so tapping into existing inventories might be helpful to not start from scratch.

Nick Coussoule: Other questions or comments? I guess, Rich, our last thing is the topics for consideration.

Denise Love: Rich, should I take this on? Just because you have been doing all the work and I have been just enjoying it.

These are our parking lot topics that we did not want to lose, but we do not really have the horsepower yet to take them on unless there is a pressing need. And the first one is the surprise medical billing act. It is included transparency and all payer claims databases, standards, and use. There is activity occurring right now. The advisory committees have been appointed through the Department of Labor. I think GAO is about to release their advisory committee. There are two of them.

But we do not really – are not taking that on unless there is some regulatory action or something that is moving through HHS in the department. It is just on our watch list. Is that a fair characterization, Rich? We will be monitoring and watching. I am tangentially involved with some of this. I will just watch it and see where this goes.

The second one comes up every year and that is scope of HIPAA to non-covered entities such as workers comp and other kinds of business associates. That, again, is just something that floated to the top and we wanted to denote that that is something that is not a priority.

HIPAA emerging technologies for electronic data exchange. This may be where we want to learn more about how the DSMOs and how the industry and how HIPAA all fit under HL7 and FHIR and how it is accommodated and then attachments and other expanded use cases. I say public health reporting is an expanded use case. I think we are watching data exchange issues and emerging technologies. Rich, I cannot remember if we are going to incorporate those into an industry-listening session later on with FHIR and HL7 or if that is just something on our maybe list.

Rich Landen: All of these items actually have a tentacle or two into this space of the Convergence 2.0. Convergence 2.0 is thinking of the whole ecosystem, not just HIPAA – clinical transactions narrowly. Whether we are talking about APCDs or the emerging technologies or some of the modernization stuff, it has impact on our view of the ecosystem. Why we were not getting these as its own project, there are aspects of it that will come under the conversation and as Denise alluded to, may be part of the scope of the listening sessions.

Denise Love: And then, again, we will be tracking the modernization. There is not direct role, but again, there are pieces of it that definitely touch on standards and social determinants. These are things we do not want to lose sight of and we will be monitoring.

Nick Coussoule: Other comments? We are right about at time for our public comment period, Rebecca. I think – we will call that section completed.

Public Comment

Rebecca Hines: Great. For those who are still in audience, if you would like to make a public comment, the raise hand feature is being enabled. You can raise your hand to have your audio unmuted. You can also send an email to ncvhsmail@cdc.gov and you can also put a comment in the Q&A that can be read into the record. Let us give that a moment. You have about four ways to do it in today's environment. If we were in the room, you would stand up and walk to the mike, but we are not in the room.

Greg Richards: We have a comment a public comment from – excuse me if I mispronounce your name – Pini Herman, please introduce yourself and state your organization.

Pini Herman: My name is Pini Herman. I am from Los Angeles. I belong to an organization called (indiscernible). We have been having quite a bit of trouble getting comparable data to evaluate our effectiveness of going out to impacted communities and arranging vaccinations and then comparing them to data from our local health department.

The primary issue is that it is not being geocoded into anything that we can compare, for example, on the census track level. I was just wondering whether CDC made this a requirement or there was an upstream requirement for data coding. I am not sure the data arrives above the state level with addresses. But this is a real problem, destiny by zip code is known, but zip codes – constructs – just much too large to evaluate.

Rebecca Hines: Thank you for your comment.

Anyone else, Greg?

Greg Richards: No one so far.

Rebecca Hines: I am going to check the email account. There is no email.

Greg, anything else coming in?

Greg Richards: No one has their hands raised. There were a few questions in the Q&A box, but I believe they were mostly all earlier in the day.

Rebecca Hines: Right and we will append the Q&A to the meeting summary. I think with that, the public comment period is over.

Nick Coussoule: Let me try to bring us to a close. I think we have had a very busy and very productive and very informative and challenging meeting. I want to thank everybody who was involved in putting that together. I will start with our support teams led by obviously Rebecca, Marietta, Geneva, and all the help in the back end, our Subcommittee staff, Lorraine and Rachel and others that helped along that

way. Sharon, for good guidance in helping us and really guide us to put together that panel, which again I would really thank and Denise for coordinating that and working with that as well.

All the committee members, thank you very much for your time and attention. We know that this is a difficult thing for us sometimes to make that time, but it is important. I think we are tackling touch topics. I really appreciate your input and guidance and feedback.

The only thing I would wrap up also with is we will continue to work on the 14th Report. We will take another issue or another iteration of that. We will get another iteration of that together and out to everybody. Sorry. I forgot Maya on that as well. I am always hesitant to bring up names to thank people for their work effort because I always forget somebody. Maya is sitting right in front of me on the screen, so I apologize if I missed you.

We will get that out to everybody. It will, again, I think Rebecca indicated that the minutes from the meeting and we will get the reporting together in due course, as we usually do to make that available. I think all the presentations either are or will be on the site as soon as we go through the compliance criteria and get that set up.

But until then, again, thank everybody for your time and attention as well as our attendees outside. I think we had well over 100 today. That may be a record for us as far as the number of people interested in what we are doing.

Again, thank you all. We are formally adjourned.

(Meeting adjourned)