MEETING

STATE OF CALIFORNIA

HEALTH AND HUMAN SERVICES AGENCY

CENTER FOR DATA INSIGHTS AND INNOVATION

COMMITTEE FOR THE PROTECTION OF HUMAN SUBJECTS

FRIDAY, APRIL 24, 2026

8:31 A.M.

OFFICE OF TECHNOLOGY AND SOLUTIONS INTEGRATION (OTSI)

2870 GATEWAY OAKS DRIVE, SUITE 200

SACRAMENTO, CALIFORNIA 95833

AND

ZOOM ONLINE MEETING PLATFORM

Reported by:
Peter Petty

PETER PETTY REPORTING, CER**D-493
4632 Freeman Way, Sacramento, California 95819
916-889-2803

APPEARANCES

COMMITTEE MEMBERS

Catherine Hess, PhD, Chair

Larry Dickey, MD, MPH, Vice Chair

Maria Dinis, PhD, MSW

Jonni Johnson, PhD

David Lang, PhD

Laura Lund, MA

Philip Palacio, EdD, MS

Juan Ruiz, MD, Dr.PH, MPH

John Schaeuble, PhD, MS

Lemeneh Tefera, MD, MSc

Maria I. Ventura, PhD

CPHS STAFF PRESENT

Agnieszka Rykaczewska, PhD, Administrator

Sussan Atifeh, Staff Services Analyst

Nicholas Zadrozna

ALSO, PRESENT

CDII

Agnieszka Rykaczewska, PhD, CDII Deputy Director

OTSI

Adam Dondro, CalHHS, Chief Information Officer, and OTSI Director

DHCS

Linette T. Scott, MD, MPH, Chief Data Officer, and Deputy Director of Enterprise Data and Information Management

Geno Guerra

PRINCIPAL INVESTIGATORS AND ASSOCIATE INVESTIGATORS

Dr. Lindsey Richland, UC Irvine (UCI)

Ms. Ella Rose, UCI

Ms. Jane Allen Research Triangle Institute (RTI)

Ms. Kim Hayes

Mr. Vaughn Armbrister

PUBLIC

None Present

I n d e x

VICE CHAIR DICKEY: Hi. I'm Larry Dickey. I'm Acting Chair of CPHS. This is the April 24th, 2026 meeting called to order.

And could you please, Sussan, call the roll?

MS. ATIFEH: Yes, sure.

Okay, Dr. Hess?

COMMITTEE MEMBER HESS: Present.

MS. ATIFEH: Dr. Dickey?

VICE CHAIR DICKEY: Present.

MS. ATIFEH: Dr. Dinis?

COMMITTEE MEMBER DINIS: Present.

MS. ATIFEH: Dr. Johnson?

COMMITTEE MEMBER JOHNSON: Present.

MS. ATIFEH: Dr. Lang?

COMMITTEE MEMBER LANG: Present.

MS. ATIFEH: Thank you.

COMMITTEE MEMBER LANG: Present.

MS. ATIFEH: Thank you.

MS. Lund?

COMMITTEE MEMBER LUND: Present.

MS. ATIFEH: Dr. Palacio?

COMMITTEE MEMBER PALACIO: Present.

MS. ATIFEH: Dr. Ruiz?

COMMITTEE MEMBER RUIZ: Present.

MS. ATIFEH: Dr. Schaeuble?

COMMITTEE MEMBER SCHAEUBLE: I'm here.

MS. ATIFEH: Dr. Tefera?

COMMITTEE MEMBER TEFERA: Present.

MS. ATIFEH: And Dr. Ventura?

COMMITTEE MEMBER VENTURA: Present.

MS. ATIFEH: Yeah, a quorum is established.

CHAIR HESS: So, we have a quorum present.

So, we have a few administrative things to do before we get going. And the first of these is to re-swear in our chair. And to do that we have the Director of OTSI, Adam Dondro.

MR. DONDRO: Thank you. I'm used to – I've done this. I've been sworn in a few times. I'm not sure I've ever actually sworn someone in.

(Laughter)

MR. DONDRO: So, I mean, if you remember February, thank you all for your endorsement of Dr. Hess. And shortly after that, Secretary Johnson did the formal appointment. So, now we just need this plain old thing.

In my instructions, I'm told that you're supposed to meet me by the flag.

VICE CHAIR DICKEY: Oh.

(Whereupon Dr. Hess is administered her oath of office.)

(Applause)

VICE CHAIR DICKEY: I always think the Common Rule should be part of that swearing in.

(Laughter)

MR. DONDRO: So, thank you, Dr. Hess –

CHAIR HESS: No, thank you.

MR. DONDRO: -- for being willing to do this. And thank you, again, to Dr. Dickey for filling in, in the brief interim.

VICE CHAIR DICKEY: And I'm going to turn it over to Dr. Hess.

CHAIR HESS: Okay. Where are we?

DR. RYKACZEWSKA: Do we have any comments from the board or the public?

CHAIR HESS: Yes. Do we have any comments from the board or the public on this agenda item, which is the swearing in of me?

(Laughter)

COMMITTEE MEMBER SCHAEUBLE: Congratulations.

CHAIR HESS: Thank you.

DR. RYKACZEWSKA: If there are any comments from members of the public, if you could please raise your virtual hand, acknowledging there are no members of the public in the room.

And I am not seeing any virtual hands. But I see a chat with a "congratulations Dr. Hess."

CHAIR HESS: Thank you.

DR. RYKACZEWSKA: With that, we'll close public comments.

CHAIR HESS: Okay, it looks like we're moving to agenda item E.

DR. RYKACZEWSKA: No, C first, and it's Adam.

CHAIR HESS: Oh, okay. It's back to Adam. We're going to agenda item C, which is the update on potential CPHS regulations. Back over to you.

MR. DONDRO: Awesome, thank you. So, I think I mentioned when I met most of you, I had no exposure to CPHS before it moved over with us at OTSI.

So, I've been doing a lot of learning, which I appreciate. And the work that you all do is admirable. And giving your time freely to do it is greatly appreciated.

I think when I took over one of the things I heard is there had been this kind of proposal on some regulations, looking at things that had been just sitting for a long time. And so, I took that as a commitment of mine that CPHS moving within OTSI to say, like lead, not a burr just to be sticking there.

And so, thank you to Dr. Dickey, Dr. Hess, Agnieszka for helping me kind of understand so that I could go and figure out

where these things are sitting and have those conversations.

And in doing that like it's clear to me that the role that CPHS plays in safeguarding the people of California, the people that are involved in the proposals that come here is paramount. And that the research is conducted with integrity, with respect to privacy and the individuals that are involved.

And that the perspective you all brought and the work that you put in is greatly appreciated with that.

One of the things that came out is that to do some of the things that were proposed, it would require regulations. And so, OTSI doesn't do regulations, not that I'd been involved in. So, I needed to go back and say what would the process be for doing this.

And so, in looking at that, you know, we look at the IPA, outline the specific criteria for the research that this Committee looks at in comparison to what was proposed in those regulations.

And doing those regulations requires not just drafting them. That's the easy part. But there's a whole formal process that requires extensive legal review, policy analysis, it requires an economic impact assessment. You have to do public notice with extensive comment periods, with formal responses to every comment that comes in through with that, which also have to go through our legal, and then agency's legal, and sometimes beyond that.

For those formal responses and for our work we would

probably need to bring in the legal of the departments for which the data belongs to for a lot of the work we do.

And so, long story short, it's an extremely intensive process. I talked to other directors in other departments about what that takes. Went up through the approval chains to have those conversations.

And, unfortunately, at this time it's not something that we're able to take on. We don't have the resources to do it, and we'd have to make a request for those. And as you're all very aware, we've lost the resources (indiscernible) right now within the state.

And so, I think that there's a lot of opportunity to continue these conversations about the things that we want to be able to look at, but at this time is the answer is we can't move forward with those right now.

No, it's not the answer we wanted to hear. And I know people put a lot of effort into putting those together, and a lot of very careful consideration. And so, again, not that we won't continue those conversations but my commitment when we keep going was that I wasn't going to get an answer on this. So, then it's kind of just sitting there in the background, with people wondering what's going on with that.

So, that's my update. That's where we're at right now.

Happy to answer any questions and, of course, any comments.

CHAIR HESS: We can open it up to questions or comments from the board. Any of the board members online.

COMMITTEE MEMBER LUND: I have a question. What will happen to all of the work that we did. There were several of us who were on that subcommittee, who put in a lot of time and effort.

MR. DONDRO: Absolutely. So, I mean, I look to Agnieszka a little bit. I think my personal thought on that is that, I mean, we keep it and we keep it present. And as we work forward in the reviews we do that we keep track of where those things pop up, and kind of keep a list of here's what we had proposed, here are these specific proposals that came in where there was an impact on our ability to look at things the way we wanted to. And kind of articulate those, so that when we come to a point in time where we want to re-raise this, we've gotten some tangible things to talk about and say, for instance, here are the places this has come up since we first did this work and looking at it.

It's not a great answer, right, because it requires waiting. But I think that it gives us some more, like some awareness we can keep on the topic in the meantime.

DR. RYKACZEWSKA: I would add to that, that I think one of the things that my personal reflections have allowed, one of the things that I have seen since the Committee has raised these

concerns is that there is a lot more conversation across the departments about many of these issues. Like, what are the risks when we're talking about linking data, for example, has been a conversation that has been coming up more and more frequently.

So, I think even without regulations, one of the great contributions that the Committee has brought through those conversations is raising awareness of these issues and making sure that they are in the forefront of our minds as we move forward. Not just when we're talking about external researchers, but even internally within the work of the departments really thinking mindfully about, and thoughtfully about when is it okay to link data and when should we be a little bit more careful and thoughtful about doing so.

So, I just -- I do want to acknowledge that even if there's not regulations, I think just the discussions and the awareness that the board brought to this issue is already, I think, a huge contribution to what we must be better about what we do with data.

VICE CHAIR DICKEY: I like what he said about this is not the end of the discussion.

MR. DONDRO: Absolutely.

VICE CHAIR DICKEY: I think we all acknowledge that there's an issue there.

COMMITTEE MEMBER DINIS: What I have to say is my experience with these things is nothing is ever done until a disaster happens and then, all of the sudden, resources and time, and all sorts of things become available. So, we'll wait until that time and then I'm sure something will be done.

COMMITTEE MEMBER JOHNSON: I have, I guess, maybe a strategic follow-up question. Because if we're not doing something at the agency level, it kind of kicks out then to the individual centers that are in agency maybe to kind of make their own decisions on how they're regulating their own data.

And to me, that sounds like there will be a lot of individualization. So, it might actually demand more resources of us at a later time if we were to wait. Is there plans to make sure that CPHS is involved with, you know, the centers' own individual plans for their regulations while they're kind of doing it right now?

If we are deciding to kind of wait at an agency level, are we getting involved in the center discussions?

MR. DONDRO: Interesting. Not something we've discussed, but I think you're absolutely right that that's where the responsibility falls to for those considerations.

I don't know. But let us take that back and think about what it would look like because it is something we do from an agency data perspective, try to help support the departments and the work they do. And so, maybe we can look to articulate some of that. Take the work that was done and see if there's a way we could share that out as things that we think that they ought to be considering and taking account for.

CHAIR HESS: I think it could help with consistency in data releases across.

COMMITTEE MEMBER JOHNSON: Yeah, because I think it is a growing concern and departments are trying to reconcile with the -- yeah, so we're not taking an agency stance, then both try to make a stance.

MR. DONDRO: Yeah. Yeah, I think it's a great flag. And we wouldn't need regulations for that. That's the kind of guidance we issue to departments in support of their work. So, that would be more feasible.

Thank you.

CHAIR HESS: Dr. Schaeuble?

COMMITTEE MEMBER SCHAEUBLE: I guess looking back on the history I'm maybe not terribly surprised that things have turned out the way they have. But I have to raise the concern, again based on past history as an individual reviewer, that we are likely

to be faced in the future with some circumstances in which, or at least some of us, the limitations that we would be operating under without the kinds of changes that were being proposed would result in something less than optimal or appropriate in the way of protection for a review of an individual project. And I'm just, I guess, wanting to acknowledge here that I personally don't want to be placed in a position of being expected to approve something that may really go against what I would see as a moral ethical kind of standard for an appropriate level of review.

And I do fear that, that situations could arise simply because we've come very close to it on a couple of occasions in the past.

MR. DONDRO: I understand.

VICE CHAIR DICKEY: Public comment?

CHAIR HESS: Sure. Let's open this up for comments from any members of the public who may be present.

DR. RYKACZEWSKA: If you are a member of the public and you would like to make public comment, if you could please raise your virtual hand. Acknowledging no members of the public in the room. And I see no virtual hands.

No public comment.

CHAIR HESS: Okay. Not the answer that maybe wanted, but hopefully something we can revisit in the future or have a role

through guidelines, as Dr. Johnson suggested. Yes.

MR. DONDRO: Yeah.

CHAIR HESS: Thank you for the update.

MR. DONDRO: Yeah, absolutely.

CHAIR HESS: Are we ready to move on? We're skipping over to Item E.

DR. RYKACZEWSKA: Temporarily.

CHAIR HESS: Yes, a temporary mix up today of when we'll review and approve the meeting minutes.

Do I have a motion to approve the meeting minutes from February 6, 2026 full board meeting?

DR. RYKACZEWSKA: If -- before we go there --

CHAIR HESS: Yes.

DR. RYKACZEWSKA: -- I do want to acknowledge that Dr. Schaeuble should provide feedback regarding the counts for the motion on page 11. And we are correcting the amount to now read that the total was nine, in favor was six plus the acting chair vote, zero no's and two abstentions.

CHAIR HESS: Thank you for that correction. I missed that. And I also missed opening it up for public comments.

So, is there any public comment on the February 6, 2026 agenda meeting minutes.

DR. RYKACZEWSKA: If any members of the public would like

to make public comment, please raise your virtual hand. Acknowledging no members of the public in the room.

I am seeing no virtual hands. And that concludes public comments.

CHAIR HESS: Now, can I have a motion to approve the February 6, 2026 meeting minutes.

VICE CHAIR DICKEY: I make the motion.

COMMITTEE MEMBER VENTURA: I second.

THE REPORTER: Who seconded, I'm sorry?

COMMITTEE MEMBER VENTURA: I did.

MS. ATIFEH: Dr. Dinis?

COMMITTEE MEMBER DINIS: Approve.

MS. ATIFEH: Dr. Johnson?

COMMITTEE MEMBER JOHNSON: Approve.

MS. ATIFEH: Dr. Lang?

COMMITTEE MEMBER LANG: Approve.

MS. ATIFEH: Ms. Lund?

COMMITTEE MEMBER LUND: Abstain.

MS. ATIFEH: Dr. Palacio?

COMMITTEE MEMBER PALACIO: Abstain.

MS. ATIFEH: Dr. Ruiz?

DR. RYKACZEWSKA: Dr. Ruiz, you are mute.

COMMITTEE MEMBER RUIZ: I said approve.

DR. RYKACZEWSKA: Thank you, Dr. Ruiz.

MS. ATIFEH: Dr. Schaeuble?

COMMITTEE MEMBER SCHAEUBLE: Approve.

MS. ATIFEH: Dr. Tefera?

COMMITTEE MEMBER TEFERA: Approve.

MS. ATIFEH: The motion passed.

CHAIR HESS: Okay. Did I hear a motion to approve?

MS. ATIFEH: Yes.

CHAIR HESS: Okay, sorry.

Okay, we had scheduled to take a break until 9:45, before moving on to Agenda Item C. But rather than take a break and sit around for nearly an hour, we do have the research team from new project 2026-004, which is University of California Irvine, online. They're on. So, we decided we would go ahead and discuss their project rather than taking a break if you're okay with that.

COMMITTEE MEMBER JOHNSON: I am in favor. I approve.

CHAIR HESS: Okay.

(Laughter)

CHAIR HESS: Thank you.

VICE CHAIR DICKEY: You're ready to go.

CHAIR HESS: The UCI team should be -- I believe they're on line. Dr. Richmond, would you be okay with moving forward with your project? Okay, we may be catching Dr. Richland off guard.

DR. RICHLAND: Yes.

CHAIR HESS: Perfect. Hello.

DR. RICHLAND: Thank you.

COMMITTEE MEMBER JOHNSON: Great. And thank you for helping us run along.

If you could please introduce anybody on your team and then give the Committee a short overview of your application?

DR. RICHLAND: Yeah, sure. So, my name is Dr. Lindsey Richland. I'm a Professor at the University of California at Irvine, in the School of Education. I am also the Director of the Science and Learning Lab at UCI.

And I can introduce my team. Maryam Cheraghi is my lab manager. She manages all the projects and the IRB for the team, as well as the recruitment efforts.

And Ella Rose is a student in my lab and she also is a former lab manager, so she is intimately connected with all of our recruitment efforts and all of the research that we conduct in the lab.

And I can provide a brief overview of the project, if that --

COMMITTEE MEMBER JOHNSON: Yes, that would be great.

DR. RICHLAND: Okay, sure. So, we conduct child development research at my lab and we're really interested in

understanding how children develop thinking and reasoning skills, how they develop those at home and at school. And how that informs their ability to thrive in society.

And so, one of the problems with a lot of the developmental research that's been conducted is that we have a pretty narrow view of who our children are, what their experiences are at home, and what their experience is at home, how they set them up for school.

So, you know, we sort of -- we have a vision for what makes for a successful student, but that doesn't connect with all of the (indiscernible) that are children are having.

So, the work that we're doing here is really trying to understand how children develop their thinking and memory skills. And we want to recruit others, for example children and families to participate, which is where practice comes in because it allows us to reach a broader, you know, set of families in our community.

And we're actually in a really rich diverse community here, in Orange County, and we're looking to make that connect with those families.

And then, the work that we do is behavioral in nature. We do, you know, experiments where we have children do tasks that are just fun, being like activities. They're just very low risk activities where they're problem solving, doing memories. And we

have survey, you know, questions for parents about activities in the home. But again, we're not getting into anything, you know, that would be risky for home.

And so, what we're asking here is for the birth record so that we can contact homes with families. At that point, they will either contact us, participating or not. And that's the only way we're going to use this data.

If they contact us, then they will give us the information that we need to enroll them in the study. And that includes their contact phone number, and email, but we do not require to (indiscernible). We're not going to take other information from the birth record directly into our databases.

So, yeah, with that I guess I'll turn it over and happy to answer any questions you may have.

COMMITTEE MEMBER JOHNSON: Yeah. So to, I guess, outline for the Committee the main thing that we're reviewing with this application is the researcher's request for birth and death data, primarily for recruitment.

They have separately, the specific study that they're getting recruitment permission for has already been approved by the UC Irvine IRB.

So, we were mainly looking at their recruitment materials, their proposal for linkage to prevent contacting

bereaved families. And the data they are requesting in birth and death data.

And I actually think that the study is very streamlined. I think that the research team did a really great job requesting only the minimal data necessary from birth and death data to link, to identify.

They also have a more conservative approach (indiscernible) -- risk of mortality, they are choosing not to, you know, contact those families. So, I think that that is, you know, the safest approach that they can take.

They, yeah, as Dr. Richland was describing, while they do contact individuals using CDPH birth data, they are redirected to enter or choose to enter their own information into their recruitment database. At which point it will become UC Irvine's data and no longer be related to the birth data.

So, at this time, you know, I think this study is looking very good. I will open it up to any of the Committee members for questions regarding the recruitment, or linkage procedure, or anything else.

COMMITTEE MEMBER LUND: I just wanted to say, because I think there's institutional memory on this, but we have a lot of new members, so since this happened. So, as background for this particular request, things used to come to this Committee, and if

they had ever touched state data, like the birth data, to develop a recruitment database, we would go on to review those studies even though the link has been severed with the birth data.

So, maybe seven years ago or so, all the lawyers took a look at it, our lawyers and the CDPH lawyers then went, okay, per statute the only purview we have as a Committee is when state data are involved.

So we, as a Committee, have purview over whether or not the data can be released for the recruitment process. But if they're not using any of the state data once recruitment is done, we don't have purview over the subsequent study. So that's, I think, what Dr. Johnson has here.

I just -- since we almost never get this on the record, I wanted to say it out loud so that I'm not the only person who remembers that this happened.

So, just to be clear for the Committee that this is something that we reviewed at recruitment, but not the subsequent studies that collect the data from the people who agreed to be in the study. Thank you.

COMMITTEE MEMBER VENTURA: I have a clarification question regarding the state data that you're using for recruitment. Is that being stored on a portable thumb drive? That was throughout the application that I -- the version that I read.

And I don't know if you responded to any of Dr. Johnson's questions about that. But I was just curious as to why identifiable data is being stored on portable thumb drives rather than, you know, behind university like firewalls, and protected in that way.

And if that's kept separate, obviously, from any of your analytic files on a separate thumb drive, for example.

DR. RICHLAND: Yeah, I believe Maryam can speak to this, and Ella as well. Actually, Ella unmuted, I think was part of that original decision.

When we originally went through the state, you know, we had a previous IRB and that was a request so that it would be severed from any possible connection to other online, you know, sort of potential access. Though everything we have is actually behind university firewalls and multiple levels of security.

But Ella, do you want to speak to that?

MS. ROSE: Yeah, I was going to say that the thumb drives are encrypted and they're stored within a lockbox. So, part of the reason was to be able to store it within a lockbox, within our locked office space. So, part of the reason was to have that additional level of security and additionally sort of locked up. Thanks.

DR. RICHLAND: Yeah, so it's essentially it's on a mobile device, but it's mobility is -- it's three feet from and it's kept

within their security of their building.

COMMITTEE MEMBER VENTURA: Okay.

COMMITTEE MEMBER JOHNSON: Did you have further -- okay. I believe, also, so there was a Spanish version of the recruitment information, which I believe Dr. Ruiz has already approved.

MS. ATIFEH: Yes.

COMMITTEE MEMBER JOHNSON: So, I am ready to -- this is my first non-deferred approval, so I'm ready to --

CHAIR HESS: Public comment.

COMMITTEE MEMBER JOHNSON: Oh. Ready to open up to the public.

CHAIR HESS: Are there any members of the public who would like to comment on this project?

DR. RYKACZEWSKA: If you are a member of the public, if you can raise your virtual hand at this time. Acknowledging no members of the public are in the room. And I see no virtual hands. We're ready for a motion.

COMMITTEE MEMBER JOHNSON: All right, I motion for approval, minimal risk, one year.

COMMITTEE MEMBER SCHAEUBLE: Dr. Johnson?

COMMITTEE MEMBER JOHNSON: Uh-hum?

COMMITTEE MEMBER SCHAEUBLE: Could you perhaps clarify in that motion that the Committee's action is not indicating any

opinion about the child development study, but it's only about the recruitment process?

COMMITTEE MEMBER JOHNSON: Yes. I make a motion for approval, minimum risk, one year. Given that the Committee has reviewed the research team's recruitment materials and strategy for having people participate in their data base. And was not in review of the actual study proposal that UC Irvine IRB --

VICE CHAIR DICKEY: The study procedures or whatever it is.

COMMITTEE MEMBER JOHNSON: Study procedures. What do we have here? -- people participating in their database and not of their specific study procedures that were separately reviewed by UCI's IRB.

DR. RYKACZEWSKA: Thank you.

COMMITTEE MEMBER SCHAEUBLE: Thanks.

CHAIR HESS: Do we have a second?

COMMITTEE MEMBER LUND: I'll second.

MS. ATIFEH: Okay. Dr. Dickey?

VICE CHAIR DICKEY: Approve.

MS. ATIFEH: Dr. Dinis?

COMMITTEE MEMBER DINIS: Sorry. Approve.

MS. ATIFEH: Dr. Lang?

COMMITTEE MEMBER LANG: Approve.

MS. ATIFEH: Dr. Palacio?

COMMITTEE MEMBER PALACIO: Approve.

MS. ATIFEH: Dr. Ruiz?

COMMITTEE MEMBER RUIZ: Approve.

MS. ATIFEH: Dr. Schaeuble?

COMMITTEE MEMBER SCHAEUBLE: Approve.

MS. ATIFEH: Dr. Tefera?

COMMITTEE MEMBER TEFERA: Approve.

MS. ATIFEH: And Dr. Ventura?

COMMITTEE MEMBER VENTURA: Approve.

MS. ATIFEH: Okay, the motion passed.

COMMITTEE MEMBER JOHNSON: All right. Well, thank you, Dr. Richland and team. You'll be receiving an approval letter from us that you can attach to your VSAC application and continue on with CDPH.

DR. RICHLAND: Yeah.

VICE CHAIR DICKEY: Yeah, but actually, before we go on, Laura you brought up the issue this, you know, about the data and subsequent studies.

But there’s something specific with VSAC, right, data in terms of it can’t be shared with subsequent studies, right?

COMMITTEE MEMBER LUND: Yeah, the researchers -- the person who receives the data can’t share it with other researchers.

But they can have a consortium of researchers who are part of the approval. So, she's part of the child development lab and there are several people working in that lab --

VICE CHAIR DICKEY: Right.

COMMITTEE MEMBER LUND: -- as long as those people are named in the VSAC application it can be shared amongst those folks.

VICE CHAIR DICKEY: Right. But they can have subsequent studies, but they can't share any VSAC data.

COMMITTEE MEMBER JOHNSON: Yeah, so CDPH's guidance on that was to have, you know, and kind of instead of approving the creation of a database, it would have to have a single study come in. Say that the database is -- you know, that they're recruiting for this database, if UCI chooses to have a separate study and wants to use that recruitment base, they have to come in under their approval.

MS. ATIFEH: Right.

COMMITTEE MEMBER JOHNSON: So that they can be evaluated by VSAC for their valid scientific interest.

VICE CHAIR DICKEY: Right. So, members may see some of their notices on their views about VSAC and that's where that comes from.

Do we have the researchers for the next project?

CHAIR HESS: We do. So, the next project, the

researchers from RTI should be on. This is project 2026-055. I asked -- let's make sure. Yeah, I see that they are on and Jane, are you ready to discuss the project?

DR. ALLEN: Hi. Um --

CHAIR HESS: Okay, wait, wait, wait.

VICE CHAIR DICKEY: No, wait, not yet.

CHAIR HESS: So, I have to formally recuse and leave the room for this.

VICE CHAIR DICKEY: Well, actually it's a -- so, she becomes chair and then she's got to recuse herself --

CHAIR HESS: Yeah.

VICE CHAIR DICKEY: -- so anyway.

(Laughter)

CHAIR HESS: So, wait to discuss.

VICE CHAIR DICKEY: But Dr. Hess can stay in the room if the Committee requests that she stay to answer questions, only. But she can't participate in the discussion otherwise.

So, the sense of the Committee, I would say we let her stay with that understanding that she's only going to answer questions.

CHAIR HESS: I just have to stay silent in the corner and look --

VICE CHAIR DICKEY: You have to stay silent. You can

answer questions, but you can't frown --

(Laughter)

CHAIR HESS: Yeah. I mean, I'm happy to step outside and be called in if there's any questions, as well, if that would make the board more comfortable.

COMMITTEE MEMBER SCHAEUBLE: No, not necessarily.

COMMITTEE MEMBER PALACIO: I suggest that you stay.

CHAIR HESS: Okay. Okay, I'm going to hand everything over, then, to you. And I will sit here in the corner.

VICE CHAIR DICKEY: And the primary reviewer on this project is Dr. Tefera, so take it away Dr. Tefera.

COMMITTEE MEMBER TEFERA: Sure. Thank you. Good morning, Jane. Do you want to give us an overview of the project and then I can make subsequent comments, and then bring it up for discussion with the Committee.

VICE CHAIR DICKEY: Actually, before we go further, I just want to confirm we still have a quorum.

MS. ATIFEH: Yes.

VICE CHAIR DICKEY: With her recused.

MS. ATIFEH: Without even -- you know, we need seven members present in this room. So, without Dr. Hess, we have seven.

VICE CHAIR DICKEY: So, if things change, we just need to confirm it.

Go ahead, I'm sorry.

COMMITTEE MEMBER TEFERA: You can go ahead, Jane.

DR. ALLEN: Okay. Great. Thank you very much. We are grateful for the opportunity to be here with you all today and glad to answer any questions you might have.

So, just to give you a sense or the overarching project, CDPH has implemented a media campaign to prevent and reduce cannabis use among California youth.

And this campaign has been going on for some time. There are two audiences. One, youth, and the other parents and caregivers. And so, there are advertising campaigns for each.

The campaign for youth is focused on helping youth understand that cannabis impairs memory and that it is not conducive to being used to alleviate stress.

And the parent campaign is about talking with your youth in an ongoing way about cannabis use among youth.

So, CDPH is currently or it's just now wrapping up formative research to create new advertising. And so, there's the existing campaign that's ongoing and there's a new campaign that contain advertisements that will be coming in next year.

And CDPH has contracted with us, RTI, and my colleagues Kim Hayes and Vaughn Armbrister are here with me today. CDPH has contracted with us to evaluate the campaign. And so, that is the

package that you see before you and the protocols describing our evaluation.

I wanted to start by saying our evaluation is rooted in behavior change theory. It's rooted in the evidence base on evaluating population campaigns, and our experience in this area, and also with our partnership with CDPH when they send instruments of the work plans they've created, and the certifications.

We have a framework we've developed that will guide all of our work. So, it will guide our evaluation throughout the three-year project. And there are four components of the evaluation.

The first is California and national comparison surveys. So, we will conduct 11 surveys over the evaluation period. The first evaluation will be designed to evaluate the campaign as they're currently on air.

The second and the third data collections will be baseline data collections. We'll use those to gather baseline data for the new campaign created that we'll start to see on the air next year.

And then, the subsequent ways will be time for us to document increases in awareness that we typically see as the campaign is sustained on the air. And we'll start to see, based on our experience, belief changes followed by other changes and other

outcomes, such as intention to engage in behavior and, hopefully, behavior changes.

The reason that we are including national comparison surveys is a key analysis we would like to conduct in the difference in difference analysis. It helps us to understand, in this complex environment where cannabis quality is changing so much and so rapidly, including that yesterday, from the federal government we -- it's very helpful for us to document changes in California in multiple ways. Including comparing it with what's happening in the rest of the nation, the nation as a comparison point.

I will get more into the details of that analysis. We can discuss it if you want to, but I wanted to help you understand the rationale behind that.

We'll be collecting data from youth, who are age 13 to 17. And also, parents and caregivers of youth who are ages 11 to 17. And we'll be collecting data from young adults, ages 18 through 20.

Now, for the very first data collection we're doing, we won't collect data from among all those populations since it contains less than (indiscernible) of those populations. So, the big picture, that's the plan for the three years ahead.

We will be recruiting through Dynata. They're an

established panel lender (indiscernible) -- before, and we have all the experience with protecting the privacy of the participants and security, also.

We are going to conduct three efficacy surveys. They're embedded in the surveys. So, in effect they probably have no impact on (indiscernible) -- necessarily. But the function of those is to gather, to understand the impact of short-term and media exposure to the advertising is very helpful for us to get an early look at how effective this campaign might be over time, especially in the early days of the campaign when population level awareness hasn't grown, yet.

We're going to conduct qualitative research, as you'll see in the package. The CDPH's plan is to look at the quantitative data that we collect and use those data to help us understand where they're at in our understanding that could be filled in with qualitative data.

So, as a result of the package that you're reviewing today, you'll see that the qualitative piece is under-developed. We haven't yet decided on the populations of interest. But we will submit an amendment for that work, when it comes time, and that is one year out.

And, finally, there will be a social media listening piece that doesn't relate to human subjects. It is really using

secondary data from the media vibes to understand what Californians are thinking about.

So, those are the product contours for evaluation. I'm sure you have many questions and our team is glad to answer them for you.

COMMITTEE MEMBER TEFERA: And --

DR. ALLEN: And, Dr. Tefera, if I've neglected anything important, please jump in. We were so grateful for your guidance earlier this week.

COMMITTEE MEMBER TEFERA: Yeah, thank you. Thank you for the overview.

Just to memorialize for the meeting, can you please describe a little bit more the recruitment the children and the parents, and why you differentiate 11 and 12 year olds from 13 to 17 year olds. And also, the related consent, please.

DR. ALLEN: Uh-hum, absolutely. Sure. We are -- our quantitative data collection, recruitment will be through Dynata. It's an established panel. People join this panel and know the purposes. They take surveys that are relevant to them. And, you know, as a result they gain points. The incentive is in the form of -- incentives may take other forms. They're non-monetary, they're points, or airline miles, or something like that.

So, the respondent chooses what their incentives will

look like and then they participate in these surveys.

So, it is very challenging for Dynata to -- it's very challenging for us to recruit teams. So, even though it's permissible, usually. But we are going to recruit them through their parents. We feel that's the most effective way we can reach them.

So, Dynata will reach out to people in their panel who they know are parents of young people within this age range. And they will screen the parent. And we will ask them, ascertain if they have youth in the age range. And if they do, then we will ask them to pass their device over for to take the survey.

We will only recruit 11 and 12 year olds through their parents.

Thirteen to 17 year olds will almost all be recruited through their parents. There is a small number of 13 to 17 year olds in the Dynata panel who could be recruited directly. But we don't anticipate that will be a very effective way for us to reach these, so primarily through all parents.

COMMITTEE MEMBER TEFERA: Thank you. I also noted in your submission that the only personal identifiable information will be birthdate and zip code. Can you just state for the Committee how you intend to use those identifiers and your plan to remove those identifiers, and if there are any changes to that plan

since submission.

DR. ALLEN: Sure. Thank you so much. Yes. We plan to collect birthdate and zip code. So, in the screen, early in the screener we ask the individual's age and their state of residence. And then later in the screener we ask them their birthdate and their zip code. And what we do is make sure that those data are in alignment before we screen someone in.

It's an important way for us to understand that the study participant is who they say they are, that they are the appropriate age range for the study. And also, that we're not inadvertently recruiting someone who's too young to participate in the study. And also, that they reside in the state that they are reporting.

So, those are the primary reasons that we're collecting those two forms of PII, birthdate and zip code.

However, and the only reason for collecting the zip code data and that is -- it comes a little later in analysis. We would like to do an analysis that shows campaign awareness by geographic region of the state. And zip code analysis is the best way to do that. It's a way of ensuring that people across the state are being equally served by the campaign. And feeding that information back to CDPH in case there are improvements that they can make.

Our plan is to store PII on secure servers at RTI. Only the bare minimum people who need to have access to PII on our team

will have it.

And, in addition, everybody who has access to it will have CITI training and will have -- begin a training that is required in our contract with CDPH. It's a training that is about information, privacy and security requirements. And not only will we take that training, but our data collection vendor has agreed to take that training.

And Vaughn, my colleague, Vaughn has created that training, has created a process for documenting that we -- that each of us has taken that training.

I realize, Dr. Tefera, I didn't fully answer your last question because you asked about our consenting documents. Is it -- I'd love to just tell you, give a little overview of those if that's of interest.

COMMITTEE MEMBER TEFERA: Sure. Please go ahead.

DR. ALLEN: Okay. So, we have -- oh, I'm sorry, you also asked about changes to our plan since it's been submitted, but I will get to that in a minute.

We have six consenting documents. So, we ask parents to consent to taking the parental screener. Then there is -- after they -- if the screener yields that they have a child who is eligible for the study, then we obtain parent permission for their child to participate in the study.

Then we have a consent form for the youth to take the screener. Then the youth take the screener, ascertaining that specific, you know, characteristics fit our study needs. And then, if they are eligible based on that screener, they complete a consent or an assent to complete the youth survey.

So, those are four consenting documents associated with the youth survey.

And for the adult survey, we have consent for the screener and then a second consent for the survey. Because we wanted to make sure that we're not asking any questions about consents, so we have those two levels of consent, both for the screener and the survey.

In our meeting with Dr. Tefera, he asked us what's the reading level of those documents. And we subsequently became aware that they were a little higher than we would have liked, especially for the youth consenting documents. So, we will be submitting an amendment to you all to reduce the reading level of those documents.

And coming back to the issue of PII, and date of birth and zip code, we also in our package to you indicated our plans to eliminate PII from our datasets immediately upon ascertaining that the participant is who they say they are, which would be very early in our process. So, I guess the data cleaning stage prior to

analysis.

But, subsequently, we have realized that the zip code analysis would be important to do to understand whether the campaign is reaching people across the state equally. So, we will submit an amendment to clarify that we'll be keeping that zip code data long, and so we can complete that analysis. And, in turn we will want to share that analysis with CDPH and make sure that they don't have questions about it, make sure that they don't need us to re-run anything. And once we have gone through that process, we will delete this data.

COMMITTEE MEMBER TEFERA: Thank you for those responses and detailing the consent process.

I don't have any further questions. For the Committee, I think it's a very interesting study, very challenging especially in this kind of milieu of what's happening with cannabis policy.

As far as our specific oversight of how PII is being managed, and the consent process, I think the investigators were very thoughtful in thinking about how to provide screeners for consent and then detailed consent both for the parents and the children. And I appreciate that the researchers will amend the consent forms to make them more accessible for the children as far as a reading level, and making it more understandable.

I don't have further concerns and would recommend

approval with the specific amendments to the timeframe for the --

VICE CHAIR DICKEY: Well, let's open it up to the Committee before --

COMMITTEE MEMBER TEFERA: Oh, yeah, yeah, I'm just telling you my two cents. Yeah, for the zip code changes that they described and also the amendments eventually.

And I'm happy for feedback from the Committee and the public.

VICE CHAIR DICKEY: Dr. Ventura?

COMMITTEE MEMBER VENTURA: I do have a clarifying question under study procedures. So, I understand that the Dynata will be administering -- recruiting and administering the survey, and then data will be transferred to RTI and stored on your servers.

Will the data also remain on Dynata's servers or will those be scrubbed?

DR. ALLEN: I believe those are scrubbed from the Dynata server. I'll look at either of my colleagues if they have additional information. We can certainly follow up to confirm.

COMMITTEE MEMBER VENTURA: Just curious where the data will live, because RTI is described as the primary site, but not Dynata. But if they're collecting data and storing it in any way, just curious how that would -- shoulder be included in the study

procedures.

DR. ALLEN: Right. My understanding is they don't retain any data on their site, that it comes to us as our data. I'll confirm. I'll seek to confirm right now. I'm sorry, I don't -- I believe it's true, but I would like to be able to give you a very concrete answer, and I can't do that right now. But I will find out for you.

COMMITTEE MEMBER JOHNSON: I have a question. For the zip code analysis you plan to do, so it's collected during the screener and then some people are subsequently screened out if they don't meet your eligibility criteria otherwise. Are those zip codes being kept or is it just the people's zip codes who get screened in?

DR. ALLEN: Only people who screen into the study, those are the only data that we will ever have.

COMMITTEE MEMBER JOHNSON: I see. Thank you.

DR. ALLEN: If somebody doesn't screen into the study, we will never receive their data.

VICE CHAIR DICKEY: And those are five-digit zip codes? Are those --

DR. ALLEN: Sorry?

VICE CHAIR DICKEY: Are those five-digit zip codes, is that what you're getting.

DR. ALLEN: Yes. That's right, yes.

CHAIR HESS: Dr. Schaeuble.

COMMITTEE MEMBER SCHAEUBLE: I have one particular concern about the consent documents I hope you would be able to address.

In talking about topics that would be sensitive and, therefore, potentially a risk to participants, you said there may be questions about drug usage and your other words were health behaviors, or health topics in different places that this was mentioned.

That second pairing of words seems unnecessarily obtuse and I'm particularly reflecting on your questions about racial or ethnic identity. You have, perhaps, the largest number of categories I've ever seen in a research project. It seems to me that it's up to well over 30, altogether.

And you also have very specific categories for sexual identify, as well. And it seems to me that your consent documents should really say something like drug usage and other topics such as racial and sexual identity to be a little more descriptive for people, when they're looking at these forms.

Is that something that you're able to do?

DR. ALLEN: Sure. I don't think it -- it hadn't occurred to me that those were sensitive topics. But we certainly could add

those to the -- to the consenting documents.

COMMITTEE MEMBER SCHAEUBLE: Well, it seems that some of the smaller sexual categories may well be of concern to some people. And given the -- let's face it, given the climate that we're in politically these days, some of the ethnic categories may be of real concern to individuals to reveal on a survey, also. So --

DR. ALLEN: Yeah, I see your point. So, you know that the race and ethnicity item is the new census item. And the question on sexual orientation and gender identity are those that are specified by CDPH by law, I believe.

COMMITTEE MEMBER SCHAEUBLE: Uh-hum.

DR. ALLEN: We've also been asked by the team to add a question about being intersex, which apparently is also California law to add. So, we have used the items that are -- have been, frankly, required of us. We're glad to do it. I don't mean to imply otherwise. But just to help you understand that those items were required by CDPH.

COMMITTEE MEMBER SCHAEUBLE: Not questioning the use of the items, just suggesting that the consent form should be more disclosing about potentially sensitive areas.

DR. ALLEN: Sure. We would be glad to add that.

COMMITTEE MEMBER SCHAEUBLE: One other question just to

ask you and I'm asking this partly out of interest, but also because this is obviously a very worthwhile project and one wants to have some feeling of relative certainty that the information you can acquire will actually be accurate and useful information.

Could you say a little bit about what you have been able to do to assure the quality of responses that you will obtain with the procedures that you're using. The reason I'm asking is there's been so much research I've been reading in the past few years about limitations of online survey responses, and even more particularly from panels of people who participate in the kind of thing that you're working with. Often -- often, very critical of the usefulness of the information that came from those kinds of sources.

So, obviously, you think this is workable for you and can you tell us a little bit about how you are making it workable for you?

DR. ALLEN: Yeah, thank you for that. We share your concern. You know, we too have, you know, seen -- you know, seen those critiques and had those concerns. And so, we have done, actually, a large number of things to improve the quality of our data. And I will tell you about some of them.

So, one of the things we do and we've done for many years is we have multiple attention checks within the first half of the

survey. And the attention checks are survey items. We embed them within a typical set of items that have an agree/disagree scale, which is a type of item format that is common in our surveys.

And for this item it will say, for this next question check the response strongly agree.

So, if somebody is clicking through and not paying attention and they get that answer wrong, we realize they didn't read the question.

So, we have two such items in our survey. And we do not include data for people who have missed the attention check items because we know they're not paying attention.

We also have, what I'm sure you all very familiar with, it's very standard, speech apps. Like if people go through faster than we think is reasonable for them to actually read the questions, they won't be included.

And we have a particular concern with this study because we are recruiting through parents. And we -- it has been brought to our attention that, you know, we're asking parents who provide permission to then turn their device over to their young person to complete the survey. And it's been brought to our attention that, indeed, it may be the parent who takes the survey, rather than the youth.

So, we've actually added a couple of items. And, in

fact, I don't think they're in the version that you have because we need to submit them within, and then for the survey because we added them after. And I apologize about that timing.

But these new items are designed to identify if it is a parent who's taking the survey for the youth. They're not tricky, they just ask. Are you a parent who's planning to take the survey on behalf of their youth? And if somebody says, yes, because maybe they inadvertently thought that was okay, then they're going to screen out. And they will screen out, so they won't be in our sample. We've arranged that with Dynata.

But at the end of the survey we ask another question that says, what has your experience been with this survey? And their response options are, I am young person who took the survey on my own behalf, I'm a parent who took this survey on behalf of my youth, and a parent who took the survey who supported my youth in taking the survey.

And if people say their parents took the survey on behalf of their child, we will not use that data in our analysis.

So, those are a couple of ways that we are ensuring the quality of the data.

Also, like we understand that panel data is limited in some ways, like who has joined the panel. And that is a limitation of our study. But at the same time, we are able to weight our data

to California and we have items in our survey to be able to do that. So, that will help.

COMMITTEE MEMBER SCHAEUBLE: Thank you.

DR. ALLEN: I hope I've answered some of your -- addressed some of your concerns.

COMMITTEE MEMBER SCHAEUBLE: Well, I think it's very helpful to know the kinds of things that researchers are trying to do in this regard. The research has been pretty damning in many cases and it's good to know what people are trying to do to deal with that.

DR. ALLEN: Thank you. Kim is my co-lead on this project and Vaughn is our quality. Vaughn has a couple of roles on this project, but one of them is our quality lead. So, I just want to invite Kim or Vaughn, if there's anything I forgot to mention in this part of the conversation or, indeed, at all in this discussion, please just jump in.

MS. HAYES: No, I don't have any additional comment. I think you've described our processes well.

MR. ARMBRISTER: I agree.

VICE CHAIR DICKEY: I had a couple of questions.

DR. ALLEN: Sure.

VICE CHAIR DICKEY: So, I wasn't really able to tell, the consent forms, how are they signed? Are they signed?

DR. ALLEN: It's a box. It's a checkbox at the bottom of the screen, yes, I consent, or I provide my permission, or no, I don't.

And then, we retain the data that acknowledges their consent. It represents their signature. And it's not a signature, it represents their consent to participate.

VICE CHAIR DICKEY: So, does this meet the requirements of written consent or do we need to have a waiver of written consent?

COMMITTEE MEMBER LUND: No, I think that constitutes written consent.

VICE CHAIR DICKEY: It's consent.

COMMITTEE MEMBER LUND: It's an electronic consent, which is considered to be written consent as long as they are actually acknowledging and retaining that, yeah.

VICE CHAIR DICKEY: Anybody else have any thoughts on that?

I just wanted to make sure because that's one of the things we're supposed to determine in our motion.

COMMITTEE MEMBER PALACIO: I agree. It does seem like consent and they signed.

VICE CHAIR DICKEY: Right.

COMMITTEE MEMBER JOHNSON: I have a clarifying question.

So, the application mentions the interviews and the social media listening, but today we're only approving the consent and the surveys. And the information on the interviews and social media list of name will be submitted as a future amendment. Is that correct?

DR. ALLEN: Social media listening is secondary data analysis, so we aren't collecting data for that.

COMMITTEE MEMBER VENTURA: Sorry, you are not collecting what?

DR. ALLEN: We don't collect data for social media listening. That's data that the media agency, who's created the campaign and buy the media, they will deliver those data to us. It's not data from humans, it's -- well, I pointed that it's not a primary data collection, it's a secondary data analysis.

But Kim or Vaughn, if you want to jump in because I'm feeling I might not be describing it well.

MS. HAYES: No, I think that covers it. I think it's -- yeah, we'll be -- to answer your other questions, we will be submitting an amendment for the interviews, once those plans are determined with CDPH.

VICE CHAIR DICKEY: Right.

COMMITTEE MEMBER JOHNSON: Okay, so the approval is -- I know you like lightly described the interviews here and that

they'll be conducted over Zoom, but we're not discussing, and debating, and approving that portion of the study in this application.

VICE CHAIR DICKEY: Doesn't sound like it.

COMMITTEE MEMBER JOHNSON: Okay.

VICE CHAIR DICKEY: I have another question. You know, reimbursement or, you know, for participation has been a hot topic for the Committee. In this case it could be different types of reimbursement. I've never heard of airline miles before.

But am I correct in saying that basically the reimbursement will be $75, regardless of what medium it's provided in?

DR. ALLEN: For the qualitative interviews that is different, that's not done through Dynata. We don't yet know who -- what group we'll work with, but we'll definitely work with a group. Typically, when we do qualitative interviews we -- who's the organization we work with depends on the population that we are interested in talking with.

So, for example, for Colorado the same team did some projects and at one point they want to understand the experience of LGBTQI+ young adults. And so, we worked with an organization that specialized in reaching that population.

So, for example, when CDPH tells us who they would like

to include in the qualitative piece, we'll identify an organization that's suitable for recruitment. This will be an organization that does this like regularly. That is, you know, something that they offer.

And then, recruitment in that case -- I mean, and incentives in that case will probably take the form of a gift card.

VICE CHAIR DICKEY: Right.

DR. ALLEN: We'll be responsible for that. So, that won't be points, or miles, or anything like that. That will be I would --

VICE CHAIR DICKEY: Like I said, you'll come back to us with an amendment as to what that will be and Dr. Tefera will bless it.

DR. ALLEN: We absolutely -- yes. We will have to because those plans are all informally there.

VICE CHAIR DICKEY: Okay.

DR. ALLEN: Those plans are also more than a year out. So, should you approve this package I mean it's -- we would -- that whole piece is unformulated and we would absolutely need to come back to you.

VICE CHAIR DICKEY: Okay.

DR. ALLEN: We will also need to come back to you with survey instruments for the new campaign because we don't yet know

what the campaign created will be, and the survey --

VICE CHAIR DICKEY: Right.

DR. ALLEN: -- needs to be tailored to it. So, we will need to come back to you multiple times for amendments --

VICE CHAIR DICKEY: Sure.

DR. ALLEN: -- over the course of the life of the project.

VICE CHAIR DICKEY: But for the Dynata part of it, the reimbursement is basically $75, it's just -- no?

DR. ALLEN: No.

VICE CHAIR DICKEY: That's the. So, they --

DR. ALLEN: Sorry, I misunderstood because it's actually quite a bit lower than that, so I thought you were talking about the qualitative piece.

For Dynata the range, and it's in the protocol that you have, it ranges from 25 cents to $4.50 per completed survey.

VICE CHAIR DICKEY: Oh, really low.

DR. ALLEN: That's the monetary value of the points.

VICE CHAIR DICKEY: And that's a -- how is that determined? I mean, how is that -- is it based on time? Go ahead.

DR. ALLEN: It is -- well, for one thing, like the respondent, when they join the panel they say like I would like to have a gift card, or I'd like to have airline miles, or points, or

whatever it may be. And that based on their decision there, it might be more on the lower side or the higher side, depending on how they decided to, you know, receive their incentive.

But that is the standard amount for a survey of our length and type with Dynata.

COMMITTEE MEMBER TEFERA: If I can help, because we had a similar conversation when we spoke about the study. What helped me kind of decipher the difference is that the vendor they're using, you know, has --

VICE CHAIR DICKEY: An existing panel.

COMMITTEE MEMBER TEFERA: -- an existing panel, existing approach, proprietary style for recruiting approach. It has this, you know, incentive level that those that are in the panel can choose. And depending on their frequency and use, that will drive how that vendor incentivizes their panel.

Researchers here are using this vendor to access the panel, so the researchers are not determinative or involved in the details of their vendor's incentive structure for recruiting folks that, you know, could be potential participants.

But yeah, it was a bit confusing to read. Thank you for answering all of those questions.

VICE CHAIR DICKEY: Okay. Well, I think we have another item coming up at 9:45.

DR. RYKACZEWSKA: We do, but I do have a quick just comments. We had a little bit of a discussion about Dynata's role in potentially storing the data. I know we will be checking whether that will be scrubbed. However, I do want to note that if they are storing the data at all, we will need a data security letter from them to ensure that their systems meet the security requirements.

So, just noting that, if there's not one already.

VICE CHAIR DICKEY: So, that would be put in the motion.

Any other questions? Open it up to the public, any questions from the public?

DR. RYKACZEWSKA: If you are a member of the public, who would like to make a public comment, please raise your virtual hand. Acknowledging no members of the public in the room.

And I see no virtual hands. Public comment is closed.

VICE CHAIR DICKEY: Okay. Hearing none, Dr. Tefera, do you want to make a motion. So, can you please tell me how to make the best motion?

VICE CHAIR DICKEY: Sure. For how long and what's the risk level.

COMMITTEE MEMBER TEFERA: The approval would probably be one year.

Motion for conditional one-year approval -- no, deferred

approval.

VICE CHAIR DICKEY: Uh-hum. With the following --

DR. RYKACZEWSKA: One year, minimal --

COMMITTEE MEMBER TEFERA: One year, minimal risk, yeah.

DR. RYKACZEWSKA: All right.

VICE CHAIR DICKEY: With the following changes to be reviewed by a subcommittee consisting of yourself.

COMMITTEE MEMBER TEFERA: To be -- and the changes are, so that the -- as described, the amendment describing the use of zip code and when the zip code data will be removed, or you can frame that, phrase that as you see fit.

Confirming for amendments to the reading levels for the adolescent consent documents.

Do we want to mention the scrubbing of that here, or somewhere we should mention it?

DR. RYKACZEWSKA: I would submit here confirmation of.

COMMITTEE MEMBER TEFERA: Sure. Confirmation that researchers, ventures, will scrub data and only researchers will be housing data.

DR. RYKACZEWSKA: And I might mention the data security letter here as well.

COMMITTEE MEMBER TEFERA: Sure. Is the data security letter required even before they scrub it or -- okay.

VICE CHAIR DICKEY: Yup.

And do we need to mention that the future plans regarding amendments for the interviews --

VICE CHAIR DICKEY: Yeah.

COMMITTEE MEMBER TEFERA: So, subsequent amendments once interview methods have been formalized and related incentive structure for participants.

DR. RYKACZEWSKA: And I think Dr. Schaeuble's comments about the consent expanded to include --

COMMITTEE MEMBER SCHAEUBLE: More or specificity about the various topics in the various consent documents.

COMMITTEE MEMBER TEFERA: Okay. How would you like that phrased then?

VICE CHAIR DICKEY: It would be broadening the language in the consent documents to include the --

COMMITTEE MEMBER SCHAEUBLE: Okay. Changing the consent documents to -- to include racial identity and sexual identity as potentially sensitive topics.

COMMITTEE MEMBER JOHNSON: And maybe be racial ethnicity data, race/ethnicity.

VICE CHAIR DICKEY: Well, it would be as topics that they'll be notified before they -- as topics to which they will be notified before they start the consent.

COMMITTEE MEMBER SCHAEUBLE: Let us know as potentially sensitive topics in the surveys.

VICE CHAIR DICKEY: Okay.

COMMITTEE MEMBER TEFERA: Okay. Okay, sounds good.

VICE CHAIR DICKEY: Okay. And do we have a second?

COMMITTEE MEMBER VENTURA: I'll second.

VICE CHAIR DICKEY: Second.

Ready to call the roll?

THE REPORTER: Who seconded, I'm sorry?

MS. ATIFEH: Dr. Ventura.

THE REPORTER: Thank you.

MS. ATIFEH: Dr. Dickey?

VICE CHAIR DICKEY: Approve.

MS. ATIFEH: Dr. Dinis?

COMMITTEE MEMBER DINIS: Approve.

MS. ATIFEH: Dr. Johnson?

COMMITTEE MEMBER JOHNSON: Approve.

MS. ATIFEH: Dr. Lang?

COMMITTEE MEMBER LANG: Approve.

MS. ATIFEH: Ms. Lund?

COMMITTEE MEMBER LUND: Abstain.

MS. ATIFEH: Dr. Palacio?

COMMITTEE MEMBER PALACIO: Approve.

MS. ATIFEH: Dr. Ruiz?

COMMITTEE MEMBER RUIZ: Approve.

MS. ATIFEH: Dr. Schaeuble?

COMMITTEE MEMBER SCHAEUBLE: Approve.

MS. ATIFEH: The motion passed.

VICE CHAIR DICKEY: Okay.

DR. ALLEN: Thanks.

VICE CHAIR DICKEY: Dr. Tefera?

COMMITTEE MEMBER TEFERA: Oh, yeah. Thank you to the research team. And the Committee staff will be contacting you with a follow-up letter for details. We appreciate your time.

DR. ALLEN: Thank you very much.

MS. HAYES: Thank you very much.

DR. ALLEN: Hope you all have a great day.

DR. RYKACZEWSKA: Thank you. And congratulations to Dr. Tefera on the review of his first project.

COMMITTEE MEMBER TEFERA: Hooray.

(Applause)

VICE CHAIR DICKEY: I turn the meeting back over to Dr. Hess, who remained amazingly neutral.

CHAIR HESS: Thank you. Okay, so we have -- I am looking at my agenda items. So, now, we're going to return to Agenda Item C [sic].

So, this is following up from the previous board meeting where board members expressed a desire to better understand the DHCS de-identification guidelines. Particularly in relation to how small cell counts are addressed and how they are ordered, and when they are suppressed.

And so, we invited Dr. Linette Scott, who is the DHCS Chief Data Officer and Deputy Director of Enterprise Data and Information Management, to give us a presentation on the DDG, and how decisions are made and calculations are made around de-identification within agency.

And so, with that, I will turn it over to Dr. Scott and thank her very much for being here today with us.

DR. SCOTT: Thank you so much for having us. Can you hear me? I just want to make sure you can hear me okay.

CHAIR HESS: We can hear you, yeah.

DR. SCOTT: Okay, perfect. And I have a couple of my colleagues here with me, Geno Guerra and Deepa Aggarwal who have both been very involved with this work. Geno will help with the presentation.

But I'll get started. So, we can go to the next slide, Geno.

So, we'll talk a little bit about the DDG, sort of how that relates to releasing data de-identified and how we assess

risk.

And go ahead and go to the next slide. So, just to give a little bit of context, we've actually had these guidelines in place, or a version thereof, for the last 12 years now. So, we've been using this process and this methodology for quite some time.

We did our initial Data De-identification Guideline for the Department of Health Care Services in 2014. We then revisited them and expanded in 2016, for the California Health and Human Services Agency.

So, in the last few years we've done an update that includes getting us to this point, which is now addition to .0 of the agency guidelines, and it's version 3 for DHCS in terms of these guidelines.

So, the reason we developed these guidelines was a need to have a consistent way of doing data de-identification across a large organization that produces a lot of public data. So, we needed to have a way that we could be consistent, be comfortable that we were releasing data publicly that was going to meet all of the requirements of the Information Practices Act of the Health Insurance Portability and Accountability Act related to public data.

And so, we've used these guidelines. And they are called guidelines because they provide a guide in terms of how we think

about data de-identification. They are not an absolutely. And so, really, the guidelines provide a way of thinking about risk, a way of putting some standard practice around identifying risk associated with datasets as we get ready to release data publicly.

So, we can go ahead and go to the next slide. So many of you, I'm sure, are very aware in the context of HIPAA, so at the Department of Health Care Services we are an inherent HIPAA-covered entity. And so, we are very much focused on the HIPAA data de-identification standard.

But whether you're doing -- whether the statute that guides you is HIPAA, or the State Information Practices Act, or other state-specific or federal statutes, statistical de-identification is the same.

The difference between the statute is what is the consequence of data being released that's not appropriately de-identified.

That being said, our de-identification guidelines do highlight some of the definitions that are a part of those. So, in the context of HIPAA there are a few methods for de-identification and for determination and safe harbor.

So, if we go to the next slide, safe harbor really is about looking for those specific identifiers that people don't usually question. Name, social security number, specific medical

record ID, things that are very specific identifiers of people.

And those are also covered under the Information Practices Act, right. So, this table is one of our figures in the guidelines that describe many identifiers, both in the context of the California IDA, as well as under HIPAA's Safe Harbor.

So, in the context of HIPAA Safe Harbor, the idea is that if you take all 18 of these items and you remove them all from the dataset, it creates a safe harbor for what's left that could then be de-identified and released publically.

A key component of that is that you have no actual knowledge that what's left could be used to identify. So, either direct awareness or reasonable professional awareness that it cannot be used.

There are some things in this context that people often say, oh, I'm going to do safe harbor, but then they start going at geography, more specific, say, than the state.

So, if we go to the next slide. These are kind of things that usually trip us up on the safe harbor, right. So, time periods more specific than the year. They want to do monthly reporting or quarterly reporting, geography more specific than the state, the county, a region, what have you.

And the way HIPAA was written is that -- well, I love lawyers, but it was maybe written by lawyers, I don't know, but

it's very specific because it says any geography smaller than the state, including the city, county, or (indiscernible) -- so that means Los Angeles County, even though Los Angeles County is larger than most states in the country, because they're doing it by county and it no longer meets safe harbor.

So, safe harbor has some weaknesses in it just because under safe harbor, Alpine County with 1,100 people, and Los Angeles County with 10 million people are treated the same.

But in any case, time reporting, so the more specific time, more specific geography.

And then, there's that wonderful catchall phrase other. Any other unique combination or code that can potentially identify an individual. So, rare diagnosis, unique population, special combinations of variables.

And this is where we then move into the use of our de-identification guidelines. So, again, HIPAA Safe Harbor, this is often what people think of when they say de-identified, remove those personal identifiers.

But we do have some specific areas that we need to have a process where we can treat all of our data analysis assessments consistently, but with (indiscernible).

So, if we go to the next slide, this is what -- sorry, one back. Yeah. So, this is where we cover context of expert

determination. So, again, in the context of HIPAA specifically, the HIPAA standard for de-identification defines safe harbor and expert determinations.

For our non-HIPAA covered programs in agency or across the state, they don't -- they're not bound by the HIPAA covered requirements for like the safe harbor and expert determination, but as you saw in an earlier slide, the Information Practices Act actually has the same kind of requirements related to personal identifiers. And it actually includes more because it includes things like financial data and other information that the HIPAA doesn't actually speak to.

So, expert determination specifically has this aspect around a person with appropriate knowledge to do the statistical review. But then, there's the second part of this which is documenting how and why you decided it was de-identified.

So, that's one of the things that specific about the HIPAA as opposed to the IPA is that documentation requirement. So, that was another reason we developed this guideline because it assists us with the documentation. Since every data product that leaves the Department of Health Care Services is covered under HIPAA and we need consistently to document that it is appropriately de-identified.

So, if we go to the next slide, this is kind of a high

level overview of sort of our steps in this process when we think about de-identification.

And so, the first set really focuses on the statistical aspects of de-identification.

The last two are our departmental processes around how we make sure we've checked all the boxes we need to check.

So, in terms of the first four sets, this really gets at the idea of having a systematic way of thinking about risk in any given data success.

And the reason I keep using this and guidelines is because the guidelines are about just that, assessing risk and then making a decision.

One of the things I do appreciate about the way the HIPAA standard was created is that it does speak to that idea of assessing risk. Because the only way you can have an absolutely, positive, no-risk scenario is you don't release any data.

And so, since that's not an option, we have then think about risk and how we balanced the different requirements in terms of transparency, and visibility, and ability to use the data with privacy, confidentiality and protecting individuals. So, it's always about solving the risk assessment.

So, in the context of our de-identification guidelines, we have this idea of thinking, first, what personal characteristics

are there? So, is it that we know the person's age, sex, gender? Is it that we know something about where they live? What kind of characteristics do we know about them?

Because, for example, some of the data we've produced is to say this many dollars was spent on diabetes in this year. And there's no information about the people receiving the benefit. In that case, great, we're done. No risk for re-identifying the individual in that. So, step one, look for personal characteristics.

Step two, take a look at the numerators and denominators. If all of the numerators, you know, the counts of whatever we're measuring are big, they're all
11 or greater, and the denominator is a good size, we use the number 20,000 because it ties back to some other HIPAA definitions, then it's big -- then, they're all big numbers, which inherently means they're all low risk of re-identification.

So, then, we can document. Yeah, we look, the data's got big numbers, low risk, good to go.

What the -- when they're small counts, when the numerator is less than 11 or it might be less than 11, when the denominator is smaller. So, we want to assess risk. We want to think about each dataset and what are the characteristics of that dataset and that product, and how might it be used.

And so, then we get into assessing risk. And this is where we -- I'll turn it over to Geno in a moment to talk about our scoring criteria.

But this is where we've developed kind of a set of -- a rubric to use to assess risk and put a score against it to say does this, you know, seem to be high risk or low risk based on the number of variables, what's included, all of those pieces.

So, again, this gets at the idea of being consistent in our approach to how we're assessing data. Is this considered high risk, then we do statistical masking, suppressing small numbers, or grouping variables or, you know, there's lots of ways of doing statistical masking.

And then, it moves into our departmental processes related to legal review, if necessary, and we describe that, and then our documentation around release, et cetera.

So, this kind of gives us a step-wide process as we think about risk, and as we think about how we assess these data, data product.

And again, we release over a thousand data products a year, so consistency and process is really important.

So, that kind of gives you some context in support of kind of the policy implementation implications and the requirements that we're trying to meet.

I'm going to turn it over to Geno to talk a little bit about how we do the actual assessments of that risk. And then, I think we should have time for questions.

So, Geno, I'll turn it over to you to do the next slide.

MR. GUERRA: Thanks, Dr. Scott.

Yeah, so I'm going to give a little bit of time, spend a little time on step three and how we transparently and very uniformly apply risk, assess risk to all of our data tables that leave our department and, you know, agency-wide as well.

So, we do this using something called an allocation scoring criteria. And it's something that we've been developing since the creation of the first DDG, you know, 12 years ago.

And this is a took that we use to assess the risk of a data table for re-identification, before we ever release it to the public.

And the way it works is it assigns an integer-based score to every variable that's included in the table. Like Dr. Scott said, the time reporting period, the geography of the table, any personal characteristics that are included.

And that really captures the specificity of the variables, the population size of the variable in question and that captures the risk that we could -- it quantifies the risk of those variables combined to create a risky dataset.

And like Dr. Scott said, we created this scoring system to be transparent and consistent across all of our data products and to be compliant with both the California IPA and with HIPAA.

So, a brief overview of how the scoring system works, itself. So, like I said each variable in the data table is assigned a score based on the granularity of the variable. For example, are you reporting age by year, are you reporting it by ten-year windows. There's different risks associated with those.

And the size of the underlying population that that variable can capture. For example, how many 17 year olds are there in California? How many people identify as Hispanic, or African American or, you know, male/female intersects, those variables.

And so, we do that and we add all of those variables up. So, basically, you can see on the left side we have an example of what our documentation would look like with actual numbers input.

So, there's three variables that are included in every single data product we release. And that's the number of events. That's the numerator that Dr. Scott was referring to. What's the actual number of your reporting.

If that number's really small, then we're assessing it as a high risk element. Like Dr. Scott said, large number we assess are inherently low risk.

What's the time reporting period? Are we reporting

annual counts? Are we reporting monthly counts? Or, for example, a point in time. Like how many people are on this waiver at this point in time. That's a very, you know, identifiable time.

And lastly, geographic range is present in any data table we release. That could be statewide. It could be county level. Zip codes, as we heard in the last presentation. We assess based on the highest risk variable of the geographic range.

For example, if it's by county, Alpine is a very high risk county because it's very small.

Lastly, we then look by all of the personal characteristics that are included in the table. Examples are age, sex, gender identity, race, ethnicity, immigration status, disability, education level, medical conditions, all of those will be usually separate scores and those will add to the risk.

And lastly, we assess the variable interaction. How many personal characteristics did we include? The more we include, we have to assume there's some sort of multiplicative effect of knowing multiple variables is going to create a more identifiable feature. More intersecting of circles.

And then, that results in a total score of the table. And we have an agency standard that a score of 13 or higher in a data table is a set that's high risk and we need to take actions to mask small counts.

So, as Dr. Scott mentioned, there's a few different ways we could mask datasets. The first, you know, some examples include increasing the granularity of our variables. Don't report age by year or don't report very, very specific race/ethnicities, group them into larger categories. And that can decrease the score of your table and prevent -- and increase the actual total score.

Or, as Dr. Scott mentioned, we can start masking small counts. And the small counts we like to mask are counts between one and ten individuals. We like to replace those, for example, an X that says this is a small value and the user can know it's small, but they can't know anything more specific than it's less than 11.

Another important feature is that not all personal characteristics are explicitly outlined in the Date De-Identification Guidelines. I think there are standard personal characteristics, but we have a generalized table for -- for example, if you're presenting data on Veterans in California, Veterans can be an identifiable feature. And so, we have a generalized table, this Table 30, that allows you to score a variable such as that. So, we know there are 1.3 million individuals who are Veterans in California. So, the score is +2 for that variable.

And we can do that for any variable as long as we use the number of Californians that that variable -- that meet that

condition. For example, number of Californians who are Veterans. And that allows us a very flexible way, but a very standardized way to assess the risk of any variable included in the dataset.

And lastly, another important caveat is that we have an outline in the Data De-Identification Guidelines that there are certain populations that we'll always suppress small amounts of, regardless of their score.

So, the first one is with -- are individuals who have a higher interaction with emergency services or with the justice system because that's publicly -- that's public information. So, individuals have a higher chance of having their names, pictures, birth dates, those types of things out in the public sphere. So, they have a higher level of public presence already.

Those include individuals with substance us disorders, serious mental illness, people involved in the child welfare system, people experiencing homelessness, victims of crime, or those without U.S. citizenship.

And another characteristic that we're always suppressing based on, you know, generalized protection based on, you know, current political crimes, et cetera. For example, individuals who are data specific for reproductive health, such as surgical termination of pregnancy or gender affirming care. We'll assess those always as high risk and always suppress small counts prior to

releasing the data. So, that is an exemption to the publication scoring criteria.

So, I want to move off through three quick examples to kind of help the Committee understand how we release data in different scenarios.

The first data is statewide birth data by maternal age. So, as we looked at that table of 18 link identifiers for HIPAA, there's none of those identifiers present in this table. All we know is that this is Cailfornia statewide, these are females and they have an age. And we don't know anything more specific than those three things, that we know their age.

So, we've assessed this as low risk, based on state power, and it would be approved as is.

So, what if we looked at the same data, annual birth data, but representative by county. So, that's clearly not state power because now we're presenting by county, not statewide.

So, then we would go into our publication scoring criteria because there are small counts present in the table. So, how would we assess it?

So, first, we'd assess there are small counts, so we score that +7 based on the table. The time is annual, so we score that +0. The geographic -- the smallest geography reported in the table is Alpine County, which has about 1,100 residents. So,

that's very small, so we score that a +7. We know, because this is pregnancy data, that this is all biological females, so +1. And we're reporting age range in two-year windows, which is also a very highly identifiable feature, so a +7. And we included two personal characteristics, sex and age range, so we assess a +2 on top of that.

That total score is 24, so that it's very clearly a high risk dataset. And we would say that the masking of all small counts, along with any counts which can be used to re-identify those, which we call complementary counts, would need to be suppressed.

So, they would suppress basically everything for Alpine and, typically with this at least two values for Amador before we release the data.

And that's, you know, that's very common. We suppress many of our data products that come out.

But let me provide you an example of a dataset that does not need suppression. So, again, we're looking at maternal -- or, we're looking at birth data annually, by calendar year, and we're presenting L.A. County, which is very large, by more broad age groups. So, the most large age group here is 15-year age group and represents race and ethnicity.

So, the work for this score, we see there are small

numbers present, so we would trigger a scoring of our data. There are small numbers of +7. It's annual, so +0 there.

L.A. County has 10 million residents, so we score that as a -5 because the geography is so large we actually say it's a risk decreaser.

Again, we know these are females, we know the age range is 15 years. And the race/ethnicity as presented, we have different race-ethnicity groups and how those are scored. The smallest risk -- the small population scores is a +2. And then, we have three included variables and there is a multiplicative effect on multiple variables, so we score that a +4.

This table only scored +11. And so, we would have this at a low-risk based table and we would say that, yes, there's a very low risk that you could identify people in these categories based on the variables presented.

So, there's a possible way that we could release data that's low risk and includes small numbers.

That completes our presentation and we are happy to take any questions from the Committee.

CHAIR HESS: Questions from Committee members? Oh, yes, go ahead.

COMMITTEE MEMBER LANG: So, one thing I am just generally curious about these guidelines, I've thought about, is

incorporation of synthetic data or differential privacy.

I think one of the things, as a researcher, I've found very challenging to deal with is complementary suppression so you can't (indiscernible) -- where you might only have one cell that you need to drop out, but that can surprisingly trigger a large number of cells downstream and
I'd just be sure an alternative, you know, have been explored or others.

DR. SCOTT: Yeah, so we've definitely run into that and we've had particular data products where programs, or stakeholders, perhaps, had wanted us to include clarifications, like Geno was talking about, you start adding variables and it increases risk, you get a lot more small numbers.

And so, we've had to back that back out to some extent and say, you know, we're just not going to be that granular with our public reporting because it's too much suppression.

I mean, with respect to your question, it's a great question, and for a research arena I think that makes a lot of sense.

We have not specifically done that because that's not what we get asked. So, as a government agency, people want us to say how many have you done. And if we were to come in and say, well, we synthetically created this view, we'd get tons of

questions on NCAQ's (phonetic) on various kinds of things in terms of trying to hide something, right.

So, the credibility of the information, the accuracy of the information is super important. WE want to make sure we're being publicly transparent as we communicate about the programs, the people we serve. And so, we have not done things with synthetic data, or those kinds of things.

So, that's why we go with the assessed risk, if it's high risk we're going to not show those values so that we're not creating a situation that people could be re-identified, as opposed to going down that avenue. And that was a conscious decision many years ago in terms of how we develop public data.

COMMITTEE MEMBER VENTURA: I have a question. Dr. Scott, there were a few exceptions for vulnerable populations, like those in the criminal justice system, those with severe mental illness. So, I'm thinking specifically for data coming from CDCR or from the Department of State Hospitals where they're automatically bumped into the high risk category or given a higher score.

How do researchers -- how should researchers, you know, working with that data look at just that automatic bump into the higher risk category? You know, we're looking -- this is statewide data, but often you can go to counties, because CDCR has, you know, facilities throughout the state, the same with state hospitals,

where also it immediately brings it down from the statewide level to more specific.

So, just how should researchers kind of apply DDG to that data that they're using, you know, for research?

DR. SCOTT: Yeah. So, and we've had lots of conversations with both of those departments, as well, as they think about how they publish. But we'd obviously defer to their specific situations.

That being said, when we think about researchers and, actually, this is a conversation we've had over the last decade as we've talked about the guidelines.

The idea is to assess risk. If it's higher risk, do not show small numbers. So, that means if you're doing research in those arenas, one, you're not going to show counts that are less than 11. You can show a zero, because a zero is a zero. But you're not going to show those counts that are between 1 and 10, or any rate percentages that are calculated based on numerator that's between that range.

So, it just means you're not going to show those values. So, you know, the other part, too, that I would just like, kind of from the research perspective, that often comes up as we're having these conversations is that we also worry about statistical stability and that (indiscernible)

-- right.

When you have those small counts, you know, whether it's -- I'll use birth data, since that's what Geno was using, right, let's say we have three pregnancies and deliveries in Alpine, but one of them was twins. So, we went from having three to four.

You know, from a percentage increase that's a huge increase. Where you have the same thing happen in a larger county it doesn't show from a statistical perspective.

So, that statistical kind of usability of the data often is very supportive of the data de-identification aspect as well.

So, for example, in the context of HEDIS, which are national standards in terms of how they measure for health plans in terms of quality measures and such, they, from a statistical perspective will say you don't include a measure if it has a denominator that's less than 30. Because you don't have a reliable calculation.

That's not a de-identification measure. But as soon as you remove all measures that have denominators less than 30, you basically have taken care of the numerator problem, as well, from a statistical de-identification perspective. And so, we get to where we need to be.

So, just kind of flagging that those things do, because if you look at this from a research perspective those are things

that are distinct, but often have similar outcomes in terms of what people are going to publish and how they're going to treat that data.

And so, just because something is higher risk doesn't mean we can't publish by county, doesn't mean we can publish by region, it just means we're not going to show those small counts that could then be used in combination with other data that would potentially identify an individual.

COMMITTEE MEMBER VENTURA: Thank you.

VICE CHAIR DICKEY: Dr. Tefera.

COMMITTEE MEMBER TEFERA: Hey, Dr. Scott, how are you. Just pure curiosity, you know, because, you know, the small count counties that you reference are generally rural, and rural counties experience many of the statistical challenges with the healthcare ecosystem, has there been any work just to maybe combine those counties so you can do analysis with regionalized or otherwise to kind of overcome the small number issue or has that not been fruitful?

DR. SCOTT: So, that is absolutely an option. Over the years, if you look across our departments as different programs, et cetera, many programs will publish things by region. And so, and even in the context of our health plans and healthcare services, we have some regional health plans to support rural areas. So, that's

always a choice is to create a region.

When we have, again, to the point that, you know, when I responded to David initially, folks feel very strongly about making sure that the data we release matches their area. So, for example, in our county behavioral health, especially mental health that's Medi-Cal based, I mean these are conversations from a decade ago.

We had conversations about, well, if we do -- if we group some of the counties, then we don't have to not show so many values. Doesn't matter. We want to see our county separately and on its own, right.

So, counties have their identity. They want to see their information specifically, which I completely understand and respect. So, we had conversations about do we group or do we not group.

And the answer from our stakeholders and the people that were going to use the information is don't group, even if it means there's small numbers that you're not going to show. It's more important to see my identity specifically than to see a larger number that's not, you know, not being shown for confidentiality reasons.

So, there's lots of reasons that people think about, whether they're going to group or not group. But it does mean that some things we think about from a de-identification perspective and

some of the reasons we've created some of the lines we have, because we assume, whether we did it or not, that certain information is going to be released by county. Or that certain information is going to be released by months, or by quarter, or what have you.

Just because there's requests or there's information that comes in, people are going to want to see that. So, we have to take that into account and assume that if you do do a regional group, assume that somewhere else it's going to be shown by county, as well. And now, I'm creating a complimentary cell issue that you're going to have to worry about.

So, those are things that do stack up and we do think about. But it's also just part of the process of assessing risk, of looking at each dataset.

And again, the use of the data and how it's going to be used, and who the audience is, is just as important to be thinking about as we do this as the actual risk assessment, and then how we pair those things together, right.

So, we need to protect, but we also need to share, and how do we find that balance.

COMMITTEE MEMBER TEFERA: Yeah, that's very helpful, especially just acknowledging the stakeholder preference, in this case at the county level, or whoever the stakeholder might be and

how their preferences might --

DR. SCOTT: Uh-hum.

COMMITTEE MEMBER TEFERA: -- not necessarily be solely focused on the data privacy issue that we're talking about.

DR. SCOTT: Uh-hum. Yeah.

VICE CHAIR DICKEY: Hi Linette. It's been a while, hasn't it.

DR. SCOTT: Yeah.

VICE CHAIR DICKEY: One of the things we're asked in this Committee is to determine whether something needs to be reviewed by the Committee.

And so, the Federal Common Rule says that de-identified data is not human subjects and the IPA says the data that cannot be linked to an individual is not subject to the IPA.

So, we're increasingly getting requests from within the agency and outside of the agency to determine that something is not research. Basically, they say, well, it's de-identified data so it's not research, human subjects research, and we'd like a determination from you that it is.

And it's a challenge for us because we have to -- I don't know whether we've actually put it into our policies and procedures, that -- and do we use what you've developed for all the departments in the agency and can it go outside of the agency. And

how successful have you been in getting it outside of the agency.

DR. SCOTT: Yeah, so, for just kind of a little bit of that history, right, we started with our first version in Health Care Services in 2014. And it -- we had a multi-departmental, so we had all departments represented, did a workgroup, had input and review from NORC to develop our first agency guidelines. In 2016 is when we had the Data De-Identification Guidelines established for the Health and Human Service Agency.

And so, it was directed to be used by all departments and offices in the agency, and has been. Over the last several years, you know, realizing the world changes around us, between AI, and all kinds of other fun things, so we did go through our process. And again, Deepa really led that process over the last three years or so, with representation from all of the departments in the agency to do these updates.

And Geno has now taken over for both. So, a huge, huge thanks to all of the team that did that.

But, so, the updates that we made to the guidelines take into account some of these things that have changed, again continuing to work on our consistency and type things.

We are using the guidelines as a way to help say this is how we expect data de-identification to occur. So, for example, at DHCS we have contracts with our counties and our plans. They don't

have to use our guidelines to be HIPAA compliant, but they do have to be HIPAA compliant, which means they have to follow safe harbor or expert determination.

It is really helpful for us to be able to point to our guidelines and say here's an example of how you can do this. You can do it other ways. You can provide us the documentation of how you're doing that, and what have you.

So, the advantage of referencing to these guidelines is it does give kind of a consistent way of thinking about it, calling out the issues, and looking at risk.

So, HCAI has incorporated the guidelines into processes and requirements related to the Health Care (indiscernible) database. And we continue to reference this in all of our data sharing agreements between state departments and other entities to say de-identification will be performed consistent with our guidelines.

And so, it has been very successful for us in terms of being a way of having the conversation.

To kind of your general concern, people will come to us and say, well, we're going to do de-identified data. And my first question is always, well, what is your definition of de-identified.

VICE CHAIR DICKEY: Right.

DR. SCOTT: Because usually it's not quite as strict as

our definition.

VICE CHAIR DICKEY: Right.

DR. SCOTT: Or it's not as well documented. So, we often then go back and forth and say, well, okay, here's how we describe it. You know, we start with the law and then we say here's how we've operationalized this. We want to see you do something similar.

So, and especially when we get into the research community that's not HIPAA grounded, I think you see a lot more definitions of what is de-identified. HIPAA is really specific. You know, limited dataset, de-identified, et cetera, with that standard.

And so, that's been in negotiations, even across our agency, right, because we have a number of departments that are not covered by HIPAA. And so, but they are covered by the Information Privacy Act. So, we kind of have been negotiating that every year amongst ourselves.

So, I appreciate and support the questions that come up when somebody says de-identified. And I will say these guidelines have been really helpful for us in having that conversation.

VICE CHAIR DICKEY: I just want to say the challenge we face as a Committee is how do we operationalize this.

DR. SCOTT: Uh-hum.

VICE CHAIR DICKEY: And do we say it's up to the departments, individually, that are releasing the data to determine that it's de-identified, and based on the criteria they want to us. And does it behoove us to look at their criteria, which seems kind of beyond our expertise. But just wondered.

COMMITTEE MEMBER LUND: I just wanted to respond to your question with a comment. Which is, gosh, it sounds like it would be really helpful if CPHS had regulations so that we could make these decisions in a documented and standardized way.

(Court Reporter pause.)

VICE CHAIR DICKEY: Okay. So, I just brought up that, you know, the IPA says cannot be linked to an individual, but it doesn't define in any way what -- you know, how you determine that.

DR. SCOTT: Yeah.

VICE CHAIR DICKEY: Have you been engaged in any discussions how that might be --

DR. SCOTT: Well, again, that's why we have the guidelines. So, HIPAA does not give you operationalizing, other than safe harbor. IPA doesn't really tell you the operationalize other than in some specific identifiers.

And so, that really is why we developed the guidelines is so we had a more specific way of operationalizing what that definition means and a way of having that consistency.

So, you know, for you all to consider, I mean if consider kind of govern under our agency, our agency has adopted specific guidelines more broadly, and so that would certainly be something that you might want to consider.

Again, it's been really helpful in terms of having that operational conversation like what do we mean, and how can we be more specific. And they are guidelines, so it doesn't mean they have to do it exactly that way. But if they do something different, they have to be able to write it down and explain it.

VICE CHAIR DICKEY: And is that something we could put in our policies and procedures just for the agency?

CHAIR HESS: For any agency data release that we review under IPA --

VICE CHAIR DICKEY: Right, right.

CHAIR HESS: -- can we require adherence to the DDG or --

COMMITTEE MEMBER LUND: We cannot mandate anything, we can make recommendations.

CHAIR HESS: Yeah, we cannot mandate. Yeah, recommend. And then if --

VICE CHAIR DICKEY: If it's not, then they need to explain.

CHAIR HESS: Then they need to provide the alternative de-identification procedures they're using.

VICE CHAIR DICKEY: Right.

DR. RYKACZEWSKA: That is my understanding. We can recommend the guidelines. And similar to what Dr. Scott was saying, if they choose to not follow that or not go with that recommendation, then they need to describe the process that they will be using and why it's equivalent or supportive. And it sounds like they would have to provide that explanation to the department, as well as to us.

DR. SCOTT: Yeah.

DR. RYKACZEWSKA: Yeah.

VICE CHAIR DICKEY: And then, outside of the agency?

CHAIR HESS: I don't think we have any purview there to require it.

VICE CHAIR DICKEY: Well, we do because the IPA is under us.

CHAIR HESS: True, yeah. Could be required -- could recommend DDG to --

DR. RYKACZEWSKA: I think recommendations are fine. But let's double check with our legal team. But I think we can recommend it.

I will also just call out that there's at least one bill that I've seen recently that actually references the CalHHS Data De-Identification Guidelines. Now, it has not passed. It's very

much still in draft form, so I don't want to suggest that it will. But I do know that it is, I would say, gaining traction or attention more broadly, as well.

COMMITTEE MEMBER JOHNSON: What is it?

DR. RYKACZEWSKA: I'd have to -- there's a few of them. I'd have to look back on that.

COMMITTEE MEMBER JOHNSON: Okay, sorry.

DR. RYKACZEWSKA: Dr. Lang, I had seen a hand earlier from you, and I just want to make sure we had addressed your question.

COMMITTEE MEMBER LANG: Just in terms of, yeah, statistical additions, I was just wondering if there are any statistical packages that our colleagues from DHCS would recommend, SBC Micro, or other products that could be complimentary suppression, I know really has some people --

DR. SCOTT: Yeah. No, I mean we don't have anything for a specific perspective. We -- you know, we use our tool sets internally that we have, but we're not recommending -- as a government entity we can say there's lots of options and we recommend explore what works best for you.

VICE CHAIR DICKEY: I think HIPAA says very small risk as determined by a statistical expert, et cetera. Have you made any -- I know you have your risk scale, but have you ever tried to

quantify the risk of what is very small?

DR. SCOTT: Well, yeah. So, I mean that aspect of what is very small, I mean that's part of the reason we have the numbers that we have. Because like our partners decided 11 as a threshold for the numerator that below which you start to get into more risk. So, also the denominator of 20,000 comes out of the HIPAA documentation.

So, that's why because of the screening things, and then the other aspects are bringing together just knowing that as you get more granular, as you get more specific if you combine variables you're increasing risk.

So, that's been the context of how we've worked on this.

DR. RYKACZEWSKA: And I do want to acknowledge that we're a little bit beyond the time that Dr. Scott, I know you had for us, and just a lot of appreciation, not to cut the conversation short but just wanted to acknowledge that you may have another meeting to get to.

CHAIR HESS: Do we open it up for questions or comments from the public, if the board has no more questions.

Oh, Dr. Schaeuble?

COMMITTEE MEMBER SCHAEUBLE: I just wanted to follow up with a thought that's been roaming around in my mind for a long time here. Projects that come to the Committee may involve not

just a release of data from an agency where we have these very carefully thought out policies, and I do appreciate all of the effort that's gone into making this process as thorough as it clearly is.

But projects we're seeing may not involve just release of data from an agency, researchers may also linking that data to other data, and the researchers themselves, as we've just acknowledged in the past few minutes, may have a very incomplete understanding of what the identification really means, and may be doing things in the total research study that just are difficult for us to actually even assess what the total impact is on the protection and privacy of the information.

One thing I'm wondering here is, is there anybody in any agency, anywhere, that does any kind of fact testing to see that the processes that are being applied to try to assure this level of protection actually show up in the product that the researchers eventually distribute to the world?

That's a very difficult kind of thing to bring up, I guess, but it seems like one would want to know at some point along the way that carefully thought out concepts are actually working in the context of what is really being done with the data. At least that's the point I wanted to bring up and ask. I don't know how one answers that but, anyway, that's the comment.

DR. SCOTT: Yeah, appreciate the concern. Staffing is such that there is not staffing assigned needed to do that for all the researchers, so we expect them to be professional and follow the rules that they agreed to, and the agreements that they agree to.

COMMITTEE MEMBER SCHAEUBLE: Is it ever done for any researchers?

DR. SCOTT: Again, researchers are expected to follow their protocols at their universities, and their various environments. We are not going to be cross-checking their work. That is their work that they -- that is separate from the department.

COMMITTEE MEMBER SCHAEUBLE: Okay.

DR. SCOTT: So, no, we're not going to do that. However, if we see that there's something that's being produced that creates some questions, and then we ask questions about it, obviously we would follow through on that and then hold them accountable.

But we do not routinely go behind and check what researchers are doing. We would only be doing that for -- in cities that are under direct contract with the department, because that would be our responsibility specifically for the entities that we directly contract with.

COMMITTEE MEMBER SCHAEUBLE: Okay.

VICE CHAIR DICKEY: I do know, I think you probably still do it, that projects that are approved by the DRC, some of them are asked to present their findings to the department.

DR. SCOTT: Uh-hum. Absolutely.

VICE CHAIR DICKEY: Which we ask people when they're closing their project to let us know of their publications.

DR. SCOTT: Uh-hum.

VICE CHAIR DICKEY: I don't know that we actually have the time, et cetera, to read them to that extent, but there is some entre we have. I mean, if we looked at their publications, we could probably see, you know, did they stay within the guidelines.

But as you said, it's a workload.

CHAIR HESS: Yeah.

DR. SCOTT: Right. And they're applying to the UC first and it's their independent body of work that they're doing research for. Again, the difference between somebody that we contract to do work for us, versus an external researcher that's requesting data to do the project, and then they own their project and they have all of the responsibility for making sure they follow the rules.

CHAIR HESS: Any final thoughts, or questions, or comments from the board?

Hearing none, are there any questions, comments or thoughts from the public?

DR. RYKACZEWSKA: Acknowledging no members of the public in the room, if there is any virtual hands, please raise your virtual hand if you would like to make a comment.

I'm seeing no virtual hands and I believe there are no members of the public online, with us right now.

CHAIR HESS: I would just like to thank Dr. Scott and the DHCS team for presenting this information to us. It's been very informative and extremely helpful. Thank you so much.

VICE CHAIR DICKEY: Thank you.

DR. SCOTT: Thank you. Have a good day.

CHAIR HESS: Okay, moving on to projects who reported adverse event and/or deviations, we have had no reports of adverse events or deviations.

So, we will move forward with the next agenda item, which is are there any Committee member questions or comments on agenda items L through P? And that would be projects requiring continuing review, amendments, projects requesting reliance agreements with other IRBs, exemptions/not research approvals, or final reports?

Hearing none, are there any public comments on agenda items L through P?

DR. RYKACZEWSKA: Acknowledging there are no members of the public in the room or online.

CHAIR HESS: Okay. Are there any public comments for

items that are not on today's agenda?

DR. RYKACZEWSKA: Again, acknowledging no members of the public in the room or online.

CHAIR HESS: Okay. In that case, I will announce that the next CPHS full board meeting will be on Friday, June 5, 2026.

Thank you all for your participation and contributions today.

If there are no further items on the agenda, which I don't believe there are, I will adjourn the meeting at 10:47 a.m.

(Thereupon, the meeting was adjourned at 10:47 a.m.)

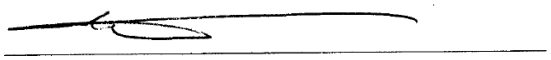
--oOo--

REPORTER'S CERTIFICATE

I do hereby certify that the testimony in the foregoing hearing was taken at the time and place therein stated; that the testimony of said witnesses were reported by me, a certified electronic court reporter and a disinterested person, and was under my supervision thereafter transcribed into typewriting.

And I further certify that I am not of counsel or attorney for either or any of the parties to said hearing nor in any way interested in the outcome of the cause named in said caption.

IN WITNESS WHEREOF, I have hereunto set my hand this 8th day of May, 2026.



PETER PETTY CER**D-493

TRANSCRIBER'S CERTIFICATE

I do hereby certify that the testimony in the foregoing hearing was taken at the time and place therein stated; that the testimony of said witnesses were transcribed by me, a certified transcriber. And I further certify that I am not of counsel or attorney for either or any of the parties to said hearing nor in any way interested in the outcome of the cause named in said caption.
IN WITNESS WHEREOF, I have hereunto set my hand this 8th day of May, 2026.

Barbara Little Certified Transcriber AAERT No. CET**D-520