[Slide 1] >> WEBINAR PRODUCER: I would now like to turn the call over to Lauren Supplee. You now have the floor. Go ahead, please.

>> LAUREN: Thank you. On behalf of the Administration for Children and Families and Health Resources Services Administration, I welcome you to the webinar. My name is Lauren Supplee. I was the Federal Project Officer on this project. And I'd like to just begin by just introducing Diane Paulsell and Sarah Avellar, two of the key staff members on this project. They'll walk you through the process that we went through and the results of the study, and then I'll rejoin you in the end to talk about some next steps. Thanks.

>> DIANE: Okay, thank you Lauren.

Before I begin, I just want to emphasize that this review was done by a large team including Sara Avellar and Brenda Jones Harden, who were the principal investigators, Patricia Del Grosso, who led the implementation [Slide 2] reviews, Emily Sama-Miller, a lead reviewer, and Barbara Gelhard, who designed the HomVEE website. And I also want to say that a – a workgroup of HHS – HHS staff contributed to all phases of the review.

So the purpose of today's webinar is to help you understand how we conducted the review and the results. I'm going to begin by explaining each step of the evidence review process. Sarah Avellar will describe in detail our criteria for assessing the quality of the studies included in the review. And then I'll present the review results, and Lauren Supplee will discuss some other details in the SIR and some next steps. At the end of the webinar, I will walk you through the HomVEE website, and then we will be available to respond to your questions.

To start off, I want to very quickly review some background on [Slide 3] the Maternal, Infant, and Early Childhood Home Visiting Program. It was established through the Patient Protection and Affordable Care Act, and it provides $1.5 billion to states over five years to establish early childhood home visiting programs. At least 75% of the funds must be used for home visiting programs with evidence of effectiveness based on well-designed and rigorous research.

[Slide 4] To inform the field and to prepare for the potential for an evidence-based home visiting initiative, in fall 2009, OPRE contracted with Mathematica Policy Research to conduct the Home Visiting Evidence of Effectiveness Review, or HomVEE as we call it for short. As I mentioned, the review was carried out under the guidance of an HHS workgroup that included staff from the Office of Planning, Research and Evaluation at ACF, the Children's Bureau at ACF, the Centers for Disease Control, [Slide 5] particularly the Division of Violence Prevention, and the National Center on Birth Defects and Developmental Disabilities, the Health Resources and Services Administration, and the Office of the Assistant Secretary for Planning and Evaluation.

For the purposes of our review, we assigned Early Childhood Home Visiting Program models very specifically. And our definition aligns with how HRSA has defined home visiting for the new federal initiatives. The models' target population must include pregnant women or families with children from birth to age five. Home visiting must be used as the primary service delivery strategy and be provided to most or all families. Home visits had to be voluntary for pregnant women, expectant fathers, and parents and caregivers of children birth to kindergarten entry.

Program models that provided services primarily in center-based settings with supplemental home visits were excluded because home visiting was not the primary service delivery strategy.

And finally, home visiting services had to be targeted to at least one of the eligible participant outcomes.

[Slide 6] Eligible participant outcomes include outcomes in these eight domains, which align with the legislation. Again, to be eligible for inclusion in the review, program models had to target at least one outcome in at least one of these domains: child health, maternal health, child development and school readiness, family economic self-sufficiency, linkages and referrals to other community services, positive

parenting practices, reduction in child maltreatment, and reduction in juvenile delinquency, family violence, and crime.

[Slide 7] So, now I'm going to begin describing the six steps we followed to conduct the review.

In step one, we sought to identify all potentially-relevant studies.

Next, we screened all of the identified studies to identify those that would be appropriate for this review.

In step three, we prioritized program models for review.

In step four, we rated the quality of the study for the prioritized program models.

In step five, we assessed the evidence of effectiveness for each model.

And in step six, we reviewed implementation information about each model.

We're going to walk you through each of these steps in more detail, but first I want to give you two important definitions.

[Slide 8] Throughout the presentation we're going to use the terms "study" and "sample." A "study" refers to a single publication or report, such as a specific journal article or a comprehensive research report on a program evaluation. The term "sample," refers to the group of children and families that participated in an evaluation of a program model and whose data were analyzed and reported together. For example, sample (inaudible) families from a particular research site or a specific cohort.

In the Home Visiting research literature, we found that often multiple studies were published about the same sample, so it's important to make sure this difference is clear.

[Slide 9] So, to identify all potentially-relevant studies, we conducted keyword searches in a wide range of research databases. And we even received some help in this from the NIH library. They were able to search additional databases that we did not have access to. All the keywords and data bases we used are listed on the HomVEE website.

We also created a custom Google search engine to search more than 50 government, research, and nonprofit websites for unpublished reports and pages, [Slide 10] sometimes referred to as the "grey literature." The list of websites we searched is also available on the HomVEE website.

We checked our search results against the bibliographies of recent literature reviews and meta analyses of Home Visiting programs, and we added relevant missing citations to our search results.

And then finally, in November 2009, we issued a call for studies and sent it to approximately 40 relevant listservs for broad dissemination.

Through this search, we identified more than 7,000 unduplicated citations, including about 150 articles submitted through the call for studies.

Our second step was to screen the 7,000-plus citations for relevance to this review. And we screened out studies for a number of reasons. We screened out studies of program models in which home visiting was not a substantial program element. For example, home visiting might have been a supplemental service or might not have been provided to most families.

We screened out studies that did not use an eligible study design. As specified in the legislation, these were randomized controlled trials and quasi-experimental designs, and Sarah is going to describe those in more detail in a few minutes.

We also retained implementation studies to collect information about how the program models were implemented. We eliminated studies in which the target population was out of range. In other words, the model did not target either pregnant women or families with children ages birth to five.

We eliminated some studies because they did not examine any child or family outcomes from the eight eligible outcome domains I reviewed earlier.

We also eliminated studies of programs that were not named programs. These were often generic home visiting interventions, and because there would be no way for a state to replicate them, weren't useful to the review.

Finally, we did not review studies published in a language other than English, and we did not review studies published before 1979. We picked 1979 as a cutoff for two reasons. First, we felt that if we looked at research that was more than 30 years old, the social context would likely be so different that findings might not be relevant enough to today's context for implementation. In addition, we wanted to make sure that we included the seminal studies of home visiting program models, and we needed to go back to 1979 to do that.

After we completed the screening, we sorted all remaining citations by program model name. At that stage, we had citations on more than 250 potential home visiting models, including about 150 with at least one randomized controlled trial or quasi-experimental design study.

[Slide 11] To rank program models for inclusion in the review, we created a point system to prioritize those with the most rigorous research evidence available. Points were assigned based on the number and design of impact studies, sample sizes of impact studies, and availability of information about implementation.

We counted the number of randomized controlled trials, single-case designs, regression discontinuity designs, and matched comparison designs for each model. To account for the greater causal validity in randomized controlled trials, single-case, and regression discontinuity designs relative to the matched comparison study, we gave them greater weight in the point system. Designs without a comparison condition did not receive any points.

Larger sample sizes are often advantageous in proving a study's ability to detect statistically-significant differences between those who participate in services and those who do not. Findings from studies using larger samples also provide more confidence that the impact will be similar as a program model is scaled up and offered to more families. Therefore, we gave greater weight to studies that used sample sizes of 50 or more families.

Implementation studies provide essential information about the requirements for launching and operating home visiting programs, such as staffing and training requirements and fidelity standards. Therefore, we gave more weight to models with available implementation information.

We eliminated programs that were implemented only in a developing world context. We also eliminated models that were no longer being implemented and for which no support for implementation is available, such as training and technical assistance, because without guidance on how to implement them, information on their effectiveness would not be useful to states.

In a final step, we compared out list of prioritized programs to an adjusted data source on the prevalence of implementation in states because we wanted to make sure that we included all of the most widely-used program models. We added one widely-implemented program model that was not on our initial prioritized list but had a sufficient amount of rigorous research. The source we used for this purpose was a 2009 Congressional Research Service report entitled Home Visitation for Families with Young Children, authored by Emilie Stoltzhaus and Karen Wrench (sp).

So this process yielded 11 early childhood home visiting program models that we included in the review. [Slide 12] And they are, listed here in alphabetical order, Early Head Start Home Visiting, [Slide 13]

Family CheckUp, Health Families America, Health Start Home Visiting, Healthy Steps, Home Instruction for Parents of Preschool Youngsters, often referred to as HIPPY, Nurse-Family Partnership, Parent-Child Home Program, Resource Mothers Program, and Safe Care.

I want to clarify that Early Head Start has three program options: a center-based option, a home-based option, and a combination option. The HomVEE review included only studies that examine the home-based option specifically. Health Start can also be implemented using a wide range of service delivery strategies. We limited our review to studies of Health Start programs that provided services primarily through home visiting.

So now I'm going to turn the presentation over to Dr. Sarah Avellar to discuss how we rated the quality of the studies we reviewed.

>> SARAH: Thanks, Diane.

So as Diane said, I am going to discuss our study rating system and some of the criteria we considered when reviewing studies. The underlying interest here is identifying effective program models, those that achieved favorable outcomes for families. Assessing whether a program model is effective requires that a study establish that a program caused the observed outcome. This may seem basic but can actually be quite challenging because it means we need to rule out all other reasons for why the outcomes might have occurred.

To link a program model with outcomes, a study attempts to establish what would have happened in the absence of program services. This is known as the counter factual. We only reviewed studies with a comparison group or condition because the comparison is designed to represent that counter factual. What would have happened to the treatment group if they had not received program services? The difference between the treatment and comparison outcomes, or the study's effects, also called impacts.

Following the legislation, we included two types of study designs, randomized controlled trials and quasi-experimental design. In randomized controlled trials, families are assigned by chance to treatment and comparison groups. The main advantage of this design is that groups are similar on average for characteristics that are known, such as race, ethnicity, or maternal education, and those that are unknown, such as parent motivation.

In terms of quasi-experimental designs, three designs were eligible for review: matched comparison group studies, single-case designs, and regression discontinuity designs. Matched comparison designs also have treatment and comparison groups, but the assignment process is not random. For example, families may volunteer for the program and those who do not want the program become the comparison group. Or a program administrator may assign families based on factors such as perceived need. This purposeful selection process has the potential to compromise the quality of the study. If the groups are different at the beginning of the study, the comparison group does not provide a good representation of the counter factual.

A limitation of the matched comparison group is that even if equivalence is established on a host of characteristics, we can never definitively determine that the groups are the same on unmeasured characteristics. I'll talk about how that was factored into our study ratings in just a minute.

Single-case and regression discontinuity designs also were eligible for the review, but these were designs not commonly used in the literature we identified. We found only two single-case designs, and both received a low rating. And we didn't identify any regression discontinuity designs. I'm not going to say too much about these designs, but we wanted to mention them because if they are well implemented, these designs have the potential to provide strong evidence of program impacts. And so these designs may become more widely used in the future.

So, across all program models, we reviewed more than 150 impact studies including randomized controlled trials [Slide 14] and quasi-experimental designs.

So this slide provides a hypothetical example for why including a comparison group is so important. in this example, we have outcomes for three different programs. The purple bars – bars represent the treatment groups and the aqua bars represent the comparison groups. So, assuming this is a favorable outcome, a higher percentage is preferable.

If we only had information from the treatment group, program three would appear to be the most effective. However, if you also consider the comparison group, represented by the aqua bars, you can see that actually program one has the largest impact because the difference between the treatment group and the comparison group outcome is the largest. Even though the absolute levels are low, treatment group members had outcomes that were twice as large at 40% than the comparison at 20%. And program three, which we had thought was the most effective, does not actually have any impact because there is no difference between the treatment and comparison groups. The sample members in program three may just have been more advantaged than those in other programs.

[Slide 15] With the HomVEE rating system, we assigned all eligible studies a rating of high, moderate, or low. The rating indicates the study's ability to produce unbiased es timates of a program model's effect. Our rating system helped distinguish between studies in which we have more confidence that the observed findings were caused by the program compared to studies in which the observed findings may be the result of other unobserved differences. For example, if the motivation of parents to seek services differed in the program and comparison conditions, then we can no longer be certain that the observed differences were because of the program.

I just want to emphasize that the study ratings do not tell us anything about the findings themselves or whether the program model is effective. For example, it's perfectly valid to have a high-quality study that shows that the program model had no effect.

[Slide 16] A high rating indicates that the study has a strong ability to estimate unbiased impacts. We have the most confidence in findings from studies with a high rating. Randomized controlled trials that are well implemented could receive a high rating. We took three factors into consideration, which I'll mention briefly now and then spend a little bit more time on each factor in just a moment.

To receive a high rating, a randomized controlled trial could not reassign families to either the treatment or comparison group after initial random assignment. This means that families could not be switched between the treatment and control groups.

High-rated studies also had to have low attrition, meaning that only a small proportion of families dropped out of the study.

And last to receive a high rating, a randomized controlled trial could not have a confounding factor. I'll give an example in just a moment, but briefly a confounding factor is an element of the design that lines up exactly with the treatment or control group. Because of this, we cannot isolate the effect of the program model from that factor.

Single-case and regression discontinuity designs also could receive a high rating if they met the Institute for Education Science's What Works Clearinghouse review standards.

[Slide 17] So in this slide, we detect a sample of interest at the top. A sample will, of course, have a range of characteristics which we represented very simply with blue people and red people. As I said, an advantage of random assignment is that the treatment and control groups will be similar, on average, as you can see at the bottom of the slide. Again, this is true whether we can measure the blue or red characteristics or whether we can't.

[Slide 18] To receive a high rating, a study could not reassign any families after random assignment was conducted. This may happen, for example, if families assigned to treatment refuse to participate or a comparison family successfully lobbies to receive treatment. Reassignment undermines the benefits of random assignment because it may create unbalanced groups for measured or unmeasured characteristics.

[Slide 19] Attrition also may compromise random assignment. As sample members are lost from the study, we may end up with a different composition in the treatment and comparison groups. This complicates our interpretation of any finding since we no longer know whether the findings are the result of the program or the different sample composition.

[Slide 20] And in this slide we've included an example of a confounding factor. One of the most common confounding factors occurs when a single home visitor provides services to all families as represented on the left panel of the slide. As you can see, in this case we cannot disentangle the effects of the program from that of the home visitor. If the study has impressive impacts, we don't know whether this was because there was a particularly skilled home visitor or because the program model itself is effective.

On the right panel of the slide, you can see an illustration of the design without a confounding factor because the program is implemented by multiple home visitors.

[Slide 21] The next rating is moderate which indicates some uncertainty about the study's ability to estimate unbiased impact. Designs that could receive a moderate rating include randomized controlled trials with some problems, such as high attrition, and quasi-experimental designs with matched comparison groups. To receive a moderate rating, both these designs had to establish the treatment and comparison groups were equivalent at the beginning of the study. For the HomVEE review, we required that baseline equivalence was established on several key characteristics: race and ethnicity, socioeconomic status, and in some cases, pre-program outcome measures. For example, if maternal employment was an outcome of interest, then a study had to establish equivalence on maternal employment at the study's onset.

You will notice that matched comparison groups could not receive a high rating. This is because of the limitation I mentioned earlier. We can never be certain that the treatment and comparison groups were equivalent on unmeasured characteristics, so [Slide 22] the highest rating a matched comparison group design can receive is moderate.

The single-case and regression discontinuity design, we rated studies as moderate if they met the What Works Clearinghouse standards with reservations.

So baseline equivalence was one of the key factors for a moderate rating. This slide provides an example of the difficulty of interpreting study findings when baseline equivalence is not established. In this example, the bars on the left represent treatment and control group outcomes before the start of program services and the bars on the right represent outcomes measured after program services ended. Again, the purple bars are treatment group outcomes and the aqua bars are comparison group outcomes. For example, percentage of employed mothers.

Outcomes measured after the program ended indicate that treatment group outcomes increased and control group outcomes decreased. However, because the groups were different at baseline, it is very difficult to figure out what this means in terms of whether the program is actually effective. In this case, the comparison group is not a good representation of what would have happened to the treatment group without the program.

[Slide 23] Finally, a low rating indicates a lack of confidence that the study can estimate unbiased impacts of the program model's effects. Any research design that we discussed could receive a low rating. A low rating was assigned if the study did not meet the standards for either high or moderate quality ratings.

And I'm going to hand it back to Diane. [Slide 24]

>> WEBINAR PRODUCER: Diane, want to make sure your phone's not on mute?

>> DIANE: Okay, thank you. Can you hear me now?

>> WEBINAR PRODUCER: Yeah. Thank you.

>> DIANE: Great.

So, after we rated all of the studies for a given program model, we used the high and moderate quality studies for that model to determine if it met the HHS criteria for an evidence-based early childhood home visiting service delivery model. According to the criteria, program models must have at least one high or moderate quality impact study of the model that found favorable, statistically-significant impacts on two or more of the eight outcome domains. Or, at least two high or moderate quality impact studies of the model that used non-overlapping analytic samples with one or more favorable, statistically-significant impacts in the same domain.

So let me give you an example. A program model meets the criteria if it has a high or moderate quality study with favorable impacts in child health and positive parenting practices. Or if it has two high or moderate quality studies based on different samples of families, both of which had favorable impacts in the child health domain.

[Slide 25] Now there are some additional requirements. First, the impacts must be found in the full sample of families. Or, if found only in a subgroup, such as families of a particular racial/ethnic group, teen parents, or another subgroup, the findings had to be replicated in two separate studies with non-overlapping samples.

Ideally, there should be a theoretical rationale for thinking that a particular program model might be effective with a specific subgroup. And researchers should plan, before an evaluation begins, to test that theory. However, we had no way to determine that whether this was the case for studies in the HomVEE review. It is possible to find subgroup impacts by chance if enough different statistical tests are performed. Moreover, researchers can keep working with the data, looking at impacts on different configurations of subgroups, until they find an impact. This is sometimes called data mining.

We wanted to have some confidence that any subgroup findings reported from the HomVEE review are likely to be real program impacts rather than chance differences. If separate findings are replicated in a second study sample, we have much more confidence that they are real program impacts.

The last requirement relates to randomized controlled trials. Following the requirements in the legislation, if the evidence of effectiveness came exclusively from randomized controlled trials, at least one statistically-significant favorable impact must be sustained for at least one year after program enrollment. And, at least one statistically-significant favorable impact must be reported in peer review journal.

Based on our individual study ratings, and assessment of the available evidence across studies for each program model, we found that seven of the 11 models we reviewed met the HHS criteria for evidence of effectiveness. [Slide 26] And they are, listed in alphabetical order: Early Head Start Home Visiting, Family CheckUp, Health Families America, Healthy Steps, Home Instruction for Parents of Preschool Youngsters, Nurse-Family Partnership, and Parents As Teachers.

[Slide 27] So in addition to assessing whether each program model met the HHS criteria for an evidence-based program, the HomVEE review looked at several other dimensions of evidence, and these are reported on the website I'm going to show you in a few minutes.

First, we looked at the quality of the outcome measures used in the studies. HomVEE differentiates outcome measures based on the likelihood that they accurately measure the outcome of interest by dividing them into primary and secondary measures. Primary measures include direct observations, such as a parent-child interaction or the home environment, direct assessment, such as assessments of children's skills and behaviors, administrative records, such as medical, school, or state records, and parent or teacher reports that are based on standardized norms measures.

Standardized measures use a uniform set of procedures for administration and scoring and use established scoring norms based on the performance of a (inaudible) example.

Secondary outcome measures are self-report measures that are not standardized.

We also looked at the duration of effects after the program ended. Many programs did not have studies that followed children and families beyond the end of the program, but for some, we were able to report on whether any favorable effects were sustained for at least one year after the end of the program.

As discussed previously, we report on whether favorable effects are replicated in the same domain in a second study using a different sample of families.

We also report on the magnitude or size of the effects when we have the information to do so. To compare magnitude across different outcomes and studies, we report what are called standard effects sizes if they are available in the studies. And we also report effects sizes that we calculated when we had the information to do so.

As discussed earlier, we report subgroup findings if they are replicated in a second study using a different research sample.

In addition to reporting favorable effects, we also report unfavorable or ambiguous effects.

During the review, we found that some effects were difficult to interpret because they could be favorable or unfavorable. An example is an increase in the number of child hospitalizations. This could indicate negative effects on child health or it could indicate increased access to needed medical care due to participation in a program, so we call those ambiguous.

We also reported all outcomes that were measured but for which no effect was found.

And finally, we report on evaluator independence by noting the funding source for each study and whether any of the study authors were also program model developers.

[Slide 28] The sixth and final step I'm going to cover is the implementation review. Our goal for the implementation review was to provide states with as much information as possible about what it takes to implement the program models we reviewed as well as information from implementation studies about practitioners' experiences implementing the models and lessons learned. To do this, we extracted implementation information from all high- and moderate-quality studies as well as standalone implementation studies of program models. And we reviewed about a hundred of these implementation studies.

We also reviewed implementation guidance and materials prepared by program developers and purveyors.

We then created detailed implementation profiles that include a description of the model, prerequisites for implementation, staff characteristics such as minimum education requirements, training requirements, availability of curricula forums and other materials, program contact information, and implementation experiences. And we gave model developers and purveyors an opportunity to review and comment on these profiles.

So I'm going to turn the presentation over to Lauren Supplee, who is going to go over some other details included in the SIR and some next steps.

>> LAUREN: Thanks Diane and Sarah.

So I'm going to go over some of the information that was listed in the SIR also related to the evidence-based programs.

The SIR released on February 8th lists the seven models determined by the HomVEE review to meet the HHS criteria of evidence of effectiveness. [Slide 29] Now as a reminder, at least 75% of the grantee's funds must be utilized for evidence-based models. The grantees can propose up to 25% of the funds for promising approaches, which includes the four models reviewed by HomVEE but were not found to currently have evidence of effectiveness that met the that met the HHS criteria.

[Slide 30] As discussed in the SIR, acceptable adaptations are changes to the model that are not tested with rigorous research but are determined by the model developers to not alter the core components related to program impact. Changes that alter the core components will not be allowed to be funded under the evidence-based funding. Any proposed adaptations, along with the information from the developers, will be reviewed and approved by HHS during the review of the state plans. Adaptations that alter the core components of a model may be eligible for funding under the promising approaches as long as they meet the other requirements stated in the SIR for promising approaches.

[Slide 31] The SIR describes four options under section three regarding the selection of the proposed home visiting models. This included two procedures. One is regarding review of the models not yet reviewed, and two is re-reviews of models already reviewed. I will discuss the requests for reviews of the additional models first.

The HomVEE project could not include review of the more than 250 home visiting models found during the literature review in the time allowed. Therefore, if a state would like to propose a model that has not yet been reviewed by the HomVEE project, the state must submit a proposal for selecting this alternate model to their HRSA project officer within 45 days of the issuance of the SIR.

[Slide 32] The evidence base [Slide 33] for the proposed model will be reviewed using the same procedures described in today's webinar. A decision will be made whether the model meets the HHS criteria for evidence of effectiveness within 45 days of receipt of the request for the review. If the model is approved, the state must provide implementation information for the approved model within 30 days of receiving the decision.

The request from a grantee to use an alternate model through their HRSA project officer must include the following: the name of the model and any other known previous names of the model. Some of the models have changed names over time. And some – and identify any affiliated organizations or researchers of the model. Provide copies of reports or journal articles the grantee may have or any known research on that model.

The proposal should also discuss how the proposed model meets the legislative requirements for the evidence-based model including: being in existence for at least three years, grounded in relevant, empirically-based knowledge, linked to program-determined outcomes, associated with a national organization or institution of higher learning, and has comprehensive home visiting program standards that ensure high quality service delivery and continuous quality improvement.

[Slide 34] The government published a Federal Register notice in the summer of 2009 requesting comments on the evidence review procedures. A number of the commenters requested the establishment of a re-review process regarding the evidence determinations. The government therefore set up a re-review process. As discussed in the SIR, it states if researchers, model developers, or others believe the application of the HHS criteria for evidence of effectiveness for a particular model contains one or more errors, and that if these errors were addressed [Slide 35] the model would meet the evidence criteria, those concerns should be submitted to the email address on the screen. Inquiries will only be accepted through this email address.

The request for re-review may be based on three factors: misapplication of the HHS criteria, missing information, and errors on the HomVEE website.

To ensure independence from the original review, the re-review team will be external to the original contractor. The re-review team will provide assurances they are free from actual or perceived conflicts of interest with that model. The re-review team will be trained and certified in the HomVEE standards. And the re-review team will use the empirical articles from the original review, any information submitted with the request for a re-review, and will make any necessary queries to the original review team.

[Slide 36] HHS will issue a final decision within 45 days of the submission of the request for the review. If a model is approved as meeting the HHS criteria for evidence of effectiveness, a state wishing to implement this model must submit an updated state plan within 30 days.

[Slide 37] All states, whether or not they made the request for review of either an additional model or a request for a re-review, will be notified of any changes to the list of the models that meet the HHS criteria.

I am now going to turn it back to Diane to discuss next steps.

>> DIANE: Thank you, Lauren.

[Slide 38] As Lauren said, we are continuing the HomVEE review in 2011. We're conducting an updated literature search on the 11 program models already included in the review as well as selecting new models to review in 2011. In addition, I'm pleased to announce [Slide 39] a new call for studies. The call is now open and will be disseminated through Listservs, and it can also be found on the HomVEE website. The call is aimed at identifying studies not previously reviewed, and the screening criteria for deciding which studies to include will be the same as those I described earlier for the first year of the review.

In addition, authors may submit new evidence or findings that build upon or expand previously-reviewed studies. But the new evidence must be presented in a new standalone paper. Submissions to the call for studies must be received no later than April 15, 2011 at the email address shown on the screen.

[Slide 40] So these are the products of the review that I'm going to show you on the website. Program Model reports present evidence of effectiveness from all studies reviewed for each program model. So there are separate reports, for example, on Family CheckUp, Healthy Families America, and so on.

Outcome Domain reports present evidence across program models for outcomes in a specific domain. For example, a state that wants to target the child health domain can look at the child health report and identify which models have favorable effects on child health outcomes and all the details about what those outcomes are, the magnitude of the effects, and the other information I described earlier.

Implementation profiles provide detailed information about what it takes to implement a program model as well as experiences documented in implementation studies. A report on the HomVEE tribal evidence review presents – presents evidence about program models implemented in tribal communities.

So now I'd like to just quickly walk you through the website.

Okay. So this is the Home page to the HomVEE website. It contains a brief description of each of the reports available, which I just described to you. The reports can be accessed through these links. They can also be accessed through the menus at the top of the screen and at the bottom. And I just want to point out in this New box, you can have a link to the HomVEE tribal evidence review and also you can access the 2011 call for studies right there.

Let's see. I want to start by showing you the About Us tab. This tab contains a project overview. It contains an executive summary that kind of summarizes the overall findings of the review. And it tells about who is involved in conducting the review.

The Review Process tab contains information about each step in the review process, basically what we've been describing during the webinar. For example, if you click on this tab, Producing Study Ratings, you'll get a detailed explanation of the study ratings and – and how we arrived at those ratings.

We also have a glossary in this tab, and it has research terms to help readers who may not be as familiar with evaluation research.

At the bottom here, we have a – a study search tool that enables you to search for studies to see if they were reviewed and to find out what the final disposition of those studies are. You can click on these boxes to pull up studies, all the studies for a given program model. You can also screen by the rating of each study. Or you can enter text into the search field, and I'm going to show you an example. I'm going to enter Duggan to see the studies that were conducted by Ann Duggan. And we'll click on this first one. It tells us that it's a Healthy Families America study. The design is correlational so it does not pass our screens because it was an ineligible study design.

I can also look for only the high-rated studies. And if I do that, could pull up another study. Also a Healthy Families America study. This one is a randomized controlled trial, and it received a high rating.

So there are 487 studies in here, and these are all the studies that we found on the 11 program models that we reviewed.

So let's look at some reports. I want to start with the Program Model reports, and we'll look at the Parents As Teachers report. This first page we call an In Brief page. It has a – a brief description of the model. It tells about the studies we reviewed. In this case it says we found 49 studies for Parents As Teachers. Sixteen of those were eligible for review.

This next section tells whether the model meets the HHS criteria. In this case, Parents As Teachers does meet the criteria. There's some additional information here about the findings of our review. And then at the bottom we provide a summary of the findings organized by each of the eight outcome domains so you can see how many outcomes were measured for each of these domains and the results. And if you want to click on these blue boxes, it provides some definitions to help you with interpreting what you're seeing in that table.

I want to point out this Study Characteristics tab. This provides some additional information about each of the studies we reviewed. And it will be helpful for states if they want to know whether, for example, a target population or community in which a study was conducted is similar to their own, they can find that information here. For example, study – under Study Participants, this tells about the sample for this particular study. It tells about the setting. This study was conducted in Cleveland, Ohio. It tells about the services that were provided, the comparison conditions, the staff characteristics of the program, and then the funding source and author affiliation at the bottom. So we have that information available for all of the high- and moderate-quality studies that we reviewed.

This next tab, Effects Shown in Research, provides information on all of the individual outcomes that were measured in the studies we reviewed. So, let's see. So, this is for a particular study. It shows all the outcomes that were reviewed. If an outcome had no effect, it is just plain text. If an outcome had a favorable effect, for example, this outcome is highlighted in green. And if it was unfavorable or ambiguous, it's highlighted in red. So you can see that.

We have a Study Ratings page, so if you want to know what was the rating for a particular study and the reason for that rating, you can look here. For example, these first two studies have a high rating. They're both randomized controlled trials. They have low attrition, no reassignment, and no confounding factors. The next group are moderate, and they're also randomized controlled trials. They had high attrition but they established baseline equivalence so they got a moderate rating. And then the last group are studies that were rated low. Typically these are matched comparison designs and randomized controlled trials that have high attrition. And typically these studies have not been able to establish baseline equivalence, and that's why they got a low rating.

And then, finally, we have a Reference list sorted by high, moderate, and low-rated studies, implementation studies, and then all the studies that did not meet the right review criteria. And you can access the Study Search utility that I showed you here as well.

Let me quickly show you the Outcome report. If you you click on one of these, you get all of the outcomes across all the models sorted for that particular outcome domain. So this is the Child Development and School Readiness report. The In Brief page here shows you a brief overview of the outcome domain. Again, the number of studies reviewed. And then it provides a summary of findings that are sorted by program model. So these are all the findings for Child Development sorted by program model.

And then the Summary of Findings provides a similar list. So this shows for Family CheckUp, all the outcomes that were measured in this domain.

If you want to know more information about the outcomes, you can go to the Summary of Outcomes Used in Research tab, and this is where we provide information about the details of the outcomes. So this is the

Child Behavior checklist. A description of the measure is provided. The data collection method, in this case, is Parent/Caregiver report. The properties of the outcome measure if they were reported by the study authors. And whether the outcome is primary or secondary.

And you can click on these little blue boxes if you want more information, for example, this provides definitions for primary and secondary outcome measures so you don't have to go to the glossary, you can just find it right here. And, similarly, this table help box provides just a bit more explanation about what's included in the table.

So the final thing I want to show you is an implementation profile. Let's look at the one for Nurse-Family Partnership. These profiles provide everything that we could find about what it takes to implement a program. So the first page is an overview. It tells about the theoretical model for the program, the target population, the outcomes, the intensity and length. It tells locations where this model has been implemented, and so forth. We have a page about the prerequisites for implementation. So you can compare to find out about the staffing requirements, whether there are requirements for home visitor education and experience, supervision requirements, ratios, and so forth.

We have some information about materials and forms that are used. And this tab also describes the fidelity standards for that model.

When it was available, we included information about costs for implementation. Typically this could be costs for training, materials, technical assistance.

There is contact information for the program model.

And then the Implementation Experiences tab summarizes the information that we gleaned from implementation studies about the experiences of practitioners. And then at the bottom, lessons learned. And, finally, this report also has a reference section that includes references for the developer recommendations that are included in the profile as well as the implementation studies that we reviewed.

So that is a quick walk through of the website. [Slide 41] So, finally, I just want to mention that if you go to the Help tab, which is all the way to the right on the website, the Contact Us page, you can submit questions directly to the website. And there is also a Frequently Asked Questions tab, and that's where responses to your questions will be posted. I know there are already some responses posted up there, so you can look at those and add to that.

And I think at this point we are concluding our presentation and we're ready to take your questions.

>> LAUREN: Thanks, Diane. There's tons of questions coming in. So I hope that we'll get to all of your questions, but if we don't, we intend to create more Frequently Asked Questions, and we'll be posting those so that you can see the answers.

So, the most frequent question I see on here is will the slides and the presentation be available after today's presentation? We are recording this presentation, and we intend to post it on the HomVEE website. We'll alert everyone with an email when that's up there so that you can view it when you need to view it.

So, I have a question for the team here. What is the racial and ethnic and income diversity of the study samples in the HomVEE review?

>> DIANE: This is Diane Paulsell. I can take that one, Lauren. Overall we found that the research samples are actually quite diverse. There are a total of 19 separate research samples across the studies that – all the studies and the models that we reviewed. And all but three of those have multiple racial/ethnic groups in them and all but one include racial/ethnic minorities. All of the samples include low-income families. So there's really a tremendous amount of diversity.

>> LAUREN: Thanks. I have another question here. How do I determine if my study was reviewed?

>> DIANE: Uh, this is Diane. I can take that one as well. The best way to determine if your study was reviewed is to pull up the Study Search tool, as I showed you, on the website, and type in your name into the open field there, and you should be able to pull up any studies on which you were an author. You could also click on the box for the program model that your study is about and pull up all the programs – the studies for a program model and see if your study is there.

>> LAUREN: Thanks. I have another question for either of you. Was the gender analyzed in any of these programs in the research?

>> DIANE: Sarah, maybe you can add on to my response. This is Diane again. The only way that would have come up, I think, is if there were subgroup analyses in a particular study that looked at gender, maybe looking at outcomes for boys versus girls in the Child Development domain or something like that. I don't think we report any subgroups by gender, so either there were none in the studies we reviewed or they didn't meet our requirement of replication.

>> SARAH: That's right. And just thinking back to the review, there were very few studies that reported effects separately by gender.

>> LAUREN: Thank you. There's a question here that I'll just answer. The individual is saying that they are having a hard time accessing the website and is it currently live. The HomVEE website is live. I think you can Google it. I think I did that this morning. The URL is also provided in the SIR, and it's also on the website or on the screen right now. So, let us know after you've accessed that if you're still not able to get on the website.

Let's see. Does the study search tool include all 7,000 citations?

>> DIANE: This is Diane again. No, it does not. It includes, I believe, 487 studies. And those are all the studies that we found for the 11 program models that we prioritized for inclusion in the review.

>> LAUREN: Thanks. Here is a two-part question that was submitted. Were maternal and infant home visiting programs included in the review, not just those about early childhood? And were birth outcomes or newborn outcomes, such as preterm births or low birth weight, included as an outcome?

>> DIANE: This is Diane again. Yes, I would say maternal and infant home visiting programs would have been included in the review. We included all program models – well, we potentially included any program model that had an outcome in one of the eight domains that I mentioned. So, Child Health is one of the domains. Maternal Health is one of the domains. So those would have been included. Now, the only question is whether they were prioritized. I don't know if you remember, we talked about a point system that we used to prioritize program models for review based on the amount of rigorous research available, whether there was implementation information available on them, and so forth. But I don't see any reason why they would not have been included in the review. And you'll have to remind me the second part of the question.

>> LAUREN: It was whether birth outcomes, newborn outcomes, such as preterm births and low birth weight, were included in outcomes.

>> DIANE: Yes. I believe you'll see some outcomes like that in the Child Health Outcome domain support.

>> LAUREN: Thank you. Another question is whether the information about implementation of these models is available to states. Does anyone want to take that?

>> DIANE: Probably the best way to get that information is to go to the Implementation tab and look at the implementation profiles for each of the program models. That's where we really summarized all the information we found about not only the requirements for implementation, but the implementation experiences of practitioners and the lessons learned that were documented in implementation studies that we reviewed.

>> LAUREN: Thanks. I have a question here that I'll take. It says, can you provide more information regarding the timing of the re-review of programs that have submitted an appeal? For example, when is the re-review expected to be complete and the results disseminated? The clock starts when a program or individual requests a re-review. And that's 45 days from the time the request was submitted to that email address that we had on the screen. It is HVEE@mathematica-mpr.com. So allow 45 days from the time the request was submitted to make a determination about the re-review. And as I mentioned earlier, all the states that are grantees will be notified if a model – if either a re-review determines that a model does meet the criteria or if we received a request for a model to be reviewed that hasn't yet been reviewed. If either of those happen and it puts a model on the list, we will let you know.

There's a follow-up question that either Diane or Sarah, if you could take this. Are there any plans to categorize research that has been processed on these models? That would be very helpful for those of us considering submissions of promising approaches.

>> DIANE: Well, this is Diane. I don't think we have comprehensive information about research that's in progress. One suggestion I would have is if there is a program model that you're considering, you know, maybe check with them about research that's in progress because we probably don't have that information.

>> LAUREN: And this is a similar question. When do you anticipate having the new literature search completed?

>> DIANE: Well, you can feel free to correct me, Sarah, if I'm wrong, but I would suspect sometime over the summer. You know, we've started the literature search. We've started reviewing studies. We have a call for studies now that's going to be open until April 15th, so it will take us some time, you know, once that closes, to sort through what we've got and complete all the reviews.

>> SARAH: That's right. Since we're also accepting admissions from states, we have to make time to review those as well.

>> LAUREN: Thank you. I have a question here. How do I know if studies were conducted in communities like the ones that I am going to be prioritizing?

>> DIANE: For that, I would really suggest that for the program models that you're considering, you go to the Program Model report and look at that Study Characteristics tab. And you can look at all the high or moderate quality studies there and look at the locations and the communities where those studies were conducted, and also get some information about the study sample. For example, was it a study conducted with teen parents, or Latino parents? Or is it a diverse group of racial/ethnic groups. You can find that information in those Study Characteristics tabs for each of the Program Model reports.

>> LAUREN: Thanks. Another question that's along the same lines. Were studies in the HomVEE review conducted in both rural and urban communities?

>> DIANE: Yes. I believe that almost all of the models have some studies that were conducted in rural communities. And, again, you can gauge that by looking at those Study Characteristics pages. Let's see, I think Early Head Start, Home Visiting, Family CheckUp, Healthy Families America, Nurse-Family Partnership, and Parents As Teachers all have at least one study sample that includes children and families from rural areas.

>> LAUREN: Great. Thanks. Did you include training or other ongoing training costs in the review?

>> DIANE: That kind of information can be found on the Implementation Profile. There is a section about training and technical assistance that will provide the information that we could find about training requirements. And there is a section on costs that provided information about fees for training. So that information is there in the Implementation Profiles.

>> LAUREN: Thanks. Are there any studies specific to American Indian and Alaskan Native populations?

>> DIANE: For that I would recommend looking at the HomVEE report on our Tribal Evidence Review, which is on the Home page of the website in that box on the left-hand side that says New. If you click on the link there, that will take you to a separate report that we did on a review of evidence of program models that were either implemented in tribal communities or the samples included substantial proportion of American Indian and Alaska Native families. And you can see that there.

>> LAUREN: Okay. Another question. Is there any consideration given to programs that may not demonstrate a large impact due to the sample being difficult to treat, such as parents with severe mental illness?

>> DIANE: Well, we did not factor the magnitude of the effects or the effect sizes into our criteria for whether the program met the criteria. We only required that the favorable impacts be statistically significant. So, the impacts could be small as long as they were statistically significant.

>> LAUREN: Thank you. Another question came in. Is there a list of models initially reviewed but that did not meet requirements to be included?

>> DIANE: Well, the 11 models that were included in the review are all listed on the website. And there is a program model report for each one. And each report states whether the model met the HHS criteria. So, yes, the programs that did not meet the HHS criteria are listed there on the website.

>> SARAH: And just to be clear, all of the information for models we reviewed are on the website. If it's not listed, we didn't review the model.

>> LAUREN: But we'll be doing ongoing review to include new models over time.

>> SARAH: That's right.

>> LAUREN: Another question. What's the evidence of effectiveness for these home visiting models for specific racial and ethnic groups?

>> DIANE: This is Diane. I can take that one. We don't have evidence, separate evidence, for particular racial/ethnic groups. We would have needed subgroup analyses for those particular racial/ethnic groups to report on that separately, and we don't have that. Furthermore, most sample sizes wouldn't have been large enough to provide sufficient statistical power to detect effects separately for the separate groups. So we can't separate out the effects for one group or another because, in a mixed sample, we don't know if the effects for one particular group could be driving the findings. But what we can say is nearly all of the samples that were used in these studies are very diverse.

>> LAUREN: Thank you. How did the review consider whether the curriculum for our models has changed over time?

>> SARAH: So, we wanted to be consistent in our review, so there had to be some continuity in the model that was being studied. So, we wanted to include only studies of a particular model. If a model has changed substantially over time, we would need new research evidence to establish the effectiveness for that particular model.

>> DIANE: One thing – this is Diane – just to add is that there are a few studies of adaptations of models. Maybe they added a new component or tested a change in the curriculum or something like that. And the comparison is between the original model and the change made to the model. And we report those findings when we have them, but we report them separately and they don't factor in to the rating of the program model in terms of whether it meets the criteria. But you will see some of those reported for some of the models.

>> LAUREN: There's just a few more questions. Are some outcomes given more weight in the review than others?

>> DIANE: No, they're not. The only requirement was that the model needed to have at least one eligible outcome. In other words, an outcome in one of the eight Outcome domains. Because states are really required to decide which outcomes they're going to target based on the needs assessment that they've done in their target communities. We felt that we would just present all the information on all the eligible outcomes and states could make those determinations as they look through the information.

>> LAUREN: Another question just came in. Is there a plan to add information pertaining to the limitations of models, such as the limited effectiveness with families experiencing domestic violence?

>> SARAH: So, that seems more like a subgroup question, so I guess I would fall back to whether there was a particular subgroup of interest, we require that the results be replicated in different samples. You can really only determine whether a model is effective if there was a separate subgroup analysis or whether the entire sample consisted of families with that characteristic. So, we're a little bit limited in what we can say based on what was presented in the research.

>> DIANE: I would just add that we do report all the findings from these studies. So there are many no effect findings. So there are lots of outcomes that were measured for which no effect was found. There are a few outcomes for which there are potentially unfavorable or ambiguous findings. And you'll see all of that if you look at the Effects Found in Research tab, and you look through the tables, you'll see, you know, the no effect, the unfavorable or ambiguous, as well as the favorable there.

>> LAUREN: Okay. Another question. Will the same criteria be used for the next round of reviews?

>> DIANE: Yes.

>> LAUREN: That was short and sweet. And a question. If the National Home Visiting Program allows each local program to determine the length of the program and the frequency of visits, how can there be fidelity to the model?

>> SARAH: So, we didn't establish standards for fidelity. That's really determined by the program developer who makes determinations of what they think are the core components, what just absolutely has to be implemented for fidelity to the model. So we just reported the information that we had, but we did not make those standards.

>> DIANE: So, those fidelity standards, if they are available, are in the Implementation Profiles. And you can look at them there. And, you know, I would just add that some of the models have a range of standards. They might not just be about the frequency and duration of the visit. But they might be standards about other dimensions of the visits and the services provided as well.

>> LAUREN: Thanks. Another question came in. What level of attrition would disqualify a study from further consideration?

>> SARAH: So there wasn't any level of attrition that would disqualify a study. It was just at what point we became concerned about potential bias because of loss of the sample. And at that point, once that threshold was crossed, the study had to establish baseline equivalence. There wasn't a point at which we would throw it out. And if you look at our website, we actually have a graph that shows exactly our cutoffs for low versus high attrition.

>> LAUREN: Okay. Thank you. There are a few questions here that we'll answer later via email or through the FAQ. They have more to do with the broad program. But I'll just give people a couple of extra minutes for any additional questions about the Evidence Review specifically.

Okay. Looks like we've answered everyone's questions sufficiently.

So, as I mentioned at the beginning of the question-and-answer period, that we are archiving this website and we will make it available on the HomVEE website and we will send you the link to where it can be found once it's up there.

Thanks to all of our presenters, and thank you to everyone for participating. Have a wonderful afternoon. Thanks.