Introducing the Home Visiting Evidence of Effectiveness (HomVEE) Version 2 Handbook: An Overview of Updated Procedures and Standards for Conducting the Annual Evidence Review

Hello everyone and thank you for attending today's event. Before we begin, we'd like to cover a few housekeeping items. At the bottom of your audience console are multiple application widgets you can use. You can expand each widget by clicking on the "Maximize" icon on the top right of the widget or by dragging the bottom right corner of the widget panel. A copy of today's slide deck is available in the Resources widget, indicated by the green file icon at the bottom of your screen.

If you have any questions for presenters, click on the Q&A widget at the bottom to submit your questions. We do capture all questions. If you have any technical difficulties, please click on the yellow Help widget. There's a question mark icon that covers common technical issues. You can also submit issues via the Q&A widget. Please note, most technical issues can be resolved by pressing F5, or Command R on Macs, to refresh our player console. Now I'll pass it off to Shirley Adelstein. Shirley, you have the floor.

Welcome!

Thank you so much. Good afternoon everyone. My name is Shirley Adelstein. I'm a senior social science research analyst in the Office of Planning, Research, and Evaluation for the Administration for Children and Families. I serve as one of the federal project officers who oversees HomVEE. We regret that Naomi Goldstein is unable to join us today, but, on behalf of Naomi and OPRE, we're delighted to welcome you, and thank you for your interest in HomVEE.

It's my pleasure to introduce Kyle Peplinski to kick off the webinar. Kyle is the Branch Chief for Policy, Data, and Technical Assistance Coordination for the Division of Home Visiting and Early Childhood Systems in the Maternal and Child Health Bureau of the Health Resources and Services Administration. HRSA is a key partner for HomVEE and oversees the administration of the Maternal, Infant, and Early Childhood Home Visiting Program in partnership with ACF. Kyle, thank you so much for being here.

Thank you, Shirley. On behalf of HRSA, I'd like to echo Shirley's thanks and welcome you to today's webinar. As with many of the research and evaluation programs related to the Maternal, Infant, and Early Childhood Home Visiting Program, or what we refer to as MIECHV, HRSA collaborates with ACF to continue to build the evidence base for home visiting and to ensure that research and evaluation activities are well-aligned with MIECHV program priorities. The HomVEE review is no exception.

Over the past two years, both HRSA and ACF have worked closely with Mathematica as they made updates to the HomVEE standards and procedures. We're pleased that, for the first time since its inception, the HomVEE review has significantly updated and published its standards and procedures through this handbook that the team will be discussing today. We hope that this handbook provides additional transparency around the review process, as well as an invaluable resource to model developers and program evaluators as they design their studies and continue to build the evidence base for home visiting. After all, the evidence base is one of the hallmarks of the MIECHV program, and the HomVEE review is key to understanding which home visiting models demonstrate evidence of effectiveness. Again, thank you for joining. And with that, I'll pass it back over to Shirley to walk us through today's event.

Agenda and Presenters

Thank you, Kyle. Our agenda for today will include content presented by HomVEE staff at OPRE and Mathematica. Broadly, the content aims to highlight updates to HomVEE standards and procedures, and

to describe plans for applying the updates, beginning with the 2021 review. I'll start us off with some brief background on the HomVEE review.

Next, Emily Sama-Miller will provide an overview of the HomVEE review process and discuss updates to procedures that guide the review. Emily is a senior researcher at Mathematica, and she serves as HomVEE's project director. Julieta Lugo-Gil will then discuss updates to the standards that guide the review. Julieta is a senior researcher at Mathematica and has overseen key portions of the updates to HomVEE standards. Finally, I'll describe HomVEE's plans for rolling out the updated Version 2 procedures and standards, and then we'll open it up to Q&A.

Introduction

HomVEE Identifies Evidence-Based Early Childhood Home Visiting Models

So, let's begin with a brief background on the HomVEE review. As home visiting models have become more widespread, understandably, there's been increased interest in offering models that have established evidence of effectiveness. In that spirit, HomVEE was launched in 2009 to review the research that examines early childhood home visiting models. HomVEE is sponsored by the Office of Planning, Research, and Evaluation for ACF in partnership with HRSA.

Now, one important use of HomVEE's results is to determine which home visiting models meet the Department of Health and Human Services criteria for an evidence-based early childhood home visiting service delivery model. This designation is a key requirement of eligibility for programs implemented with funding from the MIECHV program. In accordance with MIECHV's authorizing statute, state and territory awardees must spend the majority of their MIECHV program grants to implement evidence-based home visiting models. Up to 25 percent of funding is available to implement promising approaches that will undergo rigorous evaluation. The MIECHV is administered by HRSA in partnership with ACF.

HomVEE: A Thorough, Transparent Systematic Review

HomVEE's mission is to conduct a thorough and transparent review of early childhood home visiting models. Specifically, HomVEE provides an assessment of the evidence of effectiveness for models that serve families with pregnant women and children from birth to kindergarten entry. Of course, not all evidence is based on equally well-designed research. As a systematic review, HomVEE assesses the quality of research evidence based on our published procedures and standards. Many of those align with other federal reviews, as we'll discuss, particularly the Department of Education's influential What Works Clearinghouse.

HomVEE reviews and summarizes eligible, well-designed research on early childhood home visiting models and identifies which models are evidence-based. This determination is based on specific criteria for evidence of effectiveness established by the U.S. Department of Health and Human Services for the HomVEE review.

HomVEE Revised Procedure and Standards in 2020; Rollout in 2021

Now for the first time since launching HomVEE, ACF has substantially revised selected procedures and standards for the HomVEE review, which will be the focus for our discussion today. Now, you'll see that we'll be talking a lot about both procedures and standards, so I just want to take a moment to clarify the distinction. HomVEE procedures describe how HomVEE conducts its review process. In other words, the definitions, rules, and steps that HomVEE follows to ensure its evidence review is systematic, consistent, and free from bias. This includes how HomVEE identifies eligible research, prioritizes models for review, and reports its results. HomVEE standards describe how HomVEE assesses the rigor and quality of research, including the requirements for reported findings to receive a rating of high, moderate, or low.

As I'll be discussing in greater detail later in the webinar, the standards generally won't be retroactively. In other words, we generally won't re-review most studies with the new standards. However, we do plan to apply certain updated procedures, such as new definitions, to all research. It's important to note that the new standards aren't typically stricter. Instead, they're intended to be clearer and to better align with the latest standards from other key federal evidence reviews. Additionally, HomVEE has sought through our version two updates to address critical topics on the evolution of the home visiting field. Now I'll turn it over to Emily to provide an overview of the HomVEE review process.

Overview of the HomVEE Review Process

HomVEE is a Systematic Review

Thanks, Shirley. Before we dive into the revisions to the procedures and standards, I just wanted to take a step back and briefly summarize HomVEE's evidence review process. HomVEE uses independent staff to prioritize, review, and report on evidence of effectiveness. As a systematic review, HomVEE is structured to avoid bias. Specifically, HomVEE uses formal, search, and scoring procedures to select models and research for review; publish standards in the handbook to rate the quality of evidence of that research; and then publishes results on the HomVEE website for public viewing.

HomVEE Selects Models to Review in Two Tracks Each Year

Now, just a brief overview of the prioritization process, which we'll discuss in more detail later. Each year, HomVEE selects models to review within two tracks. Track one includes models that have not been found to be evidence-based in prior rounds of HomVEE reviews, including models that HomVEE has never reviewed before. Track two includes models that have previously been reviewed by HomVEE and have been found to be evidence-based. This process begins with identifying relevant research for review and then examining each manuscript to calculate the relative priority a model has for being reviewed by HomVEE.

HomVEE's Evidence Ratings are Low, Moderate, and High

After HomVEE selects models for review, trained reviewers apply HomVEE standards to each manuscript. Reviewers rate the findings in each manuscript and use those ratings to rate the manuscript overall. Manuscripts that receive high or moderate ratings provide evidence that at least one finding in that manuscript is attributable, at least in part, to the home visiting model examined in the study that that manuscript reports. HomVEE relies on the findings from these high and moderate rated manuscripts when applying HHS's criteria to identify evidence-based models. Those HHS criteria themselves, though, are unchanged in this update of the handbook. On the other hand, manuscripts that receive low ratings provide little evidence that the reported findings are attributable, partly or as a whole, to the home visiting model.

Models are Designated "Evidence Based" If They Meet HHS Criteria

Each year, after all manuscripts on prioritized models are reviewed, HomVEE assesses whether the model meets HHS's criteria for an evidence-based early childhood home visiting service delivery model, unless that model has already met those criteria in a previous year. HHS criteria focus on the number of favorable, statistically significant findings within and across research samples and outcome domains. Each model receives its own assessment of evidence of effectiveness.

Specifically, to be designated evidence-based, models must meet at least one of the following criteria: either at least one high or moderate rated study; shows favorable, statistically significant impacts in two or

more of the eight outcome domains; or at least two high or moderate rated studies using non-overlapping analytic study samples, show one or more favorable statistically significant impact in the same domain.

In both cases, the impacts must be found either for the full sample for the study or, if they're found for subgroups and not the full sample, they must be replicated in the same domain in two or more studies that have non-overlapping analytic study samples. Some additional criteria apply to models that have research only from a randomized control trial, and that additional information is available on the HomVEE website on the HHS criteria page, as well as in the handbook that's posted in the publications area of the website.

HomVEE Reports Effectiveness Results and Implementation Details

HomVEE reports effectiveness results and implementation details. So, the results of the research reviews go in effectiveness reports that can be accessed on the model effectiveness tab of the website pictured here, and those are organized by model. HomVEE also provides details in accompanying implementation profiles for each model about how that model is designed to be implemented. Those are also available on the website on this model implementation tab that's pictured on the right.

Review of HomVEE Procedures

Key Content

Now we'll spend some time discussing HomVEE procedures, which, as Shirley explained, is how HomVEE conducts its review. Later, we'll turn to a review of the HomVEE standards themselves. Our overview of procedures is going to highlight some key content, some new definitions of terms that were established in the version two handbook, revisions that HomVEE made to search and screening procedures, a detailed review of how models are prioritized for review. This process, the prioritization process, has had only minor revisions, but we'll go through it in detail because we know it's of high interest to HomVEE stakeholders. And then, finally, we'll give a quick overview of two other procedural changes, how HomVEE handles supplemental information and how HomVEE reports subgroup findings.

So, this presentation is going to focus both on topics that were changed and were unchanged by the version two revisions. Throughout we'll use this blue badge that you see on the bottom of your screen to denote which things that we're talking about are actually changes in the handbook. We don't have time to cover all of the nuanced changes, so we really encourage viewers of this webinar to go to the HomVEE website and view the handbook there. The first exhibit, exhibit 1.2 in that handbook, is a detailed listing of all of the changes in the handbook.

Version 2 Defines Key Terms

HomVEE defined key terms in version two of the handbook. Three of those terms are the definition of study, manuscript, and finding. So, I'll go through those now. A study is an evaluation of a home visiting model implemented with a distinct sample enrolled in the research investigation at a defined time and place by a specific researcher or research team. A study, though, might produce one or many manuscripts that describe the study results. For example, as researchers follow a sample over time, the researchers may release different manuscripts for each follow-up period.

Manuscripts may be published or unpublished research, a journal article, book chapters, working papers. Typically, one manuscript reports findings from only one study. Each manuscript usually includes multiple findings, though, and findings summarize the effect of the model on a specific study sample on specific outcome measures at a specific time point from a specific analysis.

Version 2 Revised How HomVEE Finds Relevant Research

Under version two, HomVEE made three key changes to how we find relevant research. First, HomVEE implemented a 20-year moving window for manuscripts, with an exception for manuscripts that are submitted or have previously been submitted to the annual call for research. Previously, HomVEE had searched back to 1989, which was 20 years before the review's inception.

Second, HomVEE is now going to use a modified version of the Peer Review of Electronic Search Strategies, or PRESS, method. This refines the search terms and recognizes accepted library science practice and was just an opportunity to confirm that HomVEE search terms align with the review's focus. Through this process, HomVEE refined search terms, but then HomVEE also added a few search terms based on recommendations from stakeholders and experts that were received during the public comment period in August. HomVEE also expanded the annual search to grey literature databases. These include Google Scholar, Harvard's Think Tank Search, and a pair of preprint servers, which are the Open Science Framework and medRxiv. HomVEE screens all the relevant research based on screening criteria that we cover in the handbook. These screening criteria are actually largely unchanged from the original procedures.

HomVEE Uses Prioritization Points to Select Models

Now, due to limited resources, HomVEE cannot review every model identified through the literature search each year. After identifying and screening all of the relevant research, HomVEE uses a formal process to identify which models to review that year. This process is called the prioritization process and it's guided by prioritization points, which HomVEE assigns and calculates in four steps. I'll go through those four steps now. For those viewers who are noticing that the font on the slide might be a little small to read this, it's captured from the handbook, so you can go there to see the details of the graphic. I'll go through each step individually next.

The process that we're using, beginning with the 2021 review, is largely unchanged from the process in place last year, although we did make some refinements to the point allocations based on the feedback we got during the public comment period. I'll discuss those changes a little bit later.

1. HomVEE Assigns Points to Each Eligible Manuscript

The first step is to assign points to each eligible manuscript, that is manuscripts that have passed the screening stage. HomVEE focuses on individual manuscripts rather than studies overall because one study may span years or decades, and individual manuscripts reflect the volume of new research that's being produced about a model and the current state of the evidence base.

At this stage, HomVEE assigns manuscript points based on the information in the title and abstract of the manuscript. These points emphasize well-designed impact studies as well as larger sample sizes, outcomes of interest and populations of interest that are aligned with the criteria in MIECHV's authorizing statute.

Each manuscript can earn up to 6.5 points

Each manuscript, at this stage, can earn a total of six-and-a-half points based on the six criteria shown here. For study design, manuscripts that report findings from randomized control trials, or RCTs, single case designs and regression discontinuity designs receive three points, whereas manuscripts that report findings from non-experimental comparison group design studies receive two points. For sample size, manuscripts receive one point if the total sample size is 250 or greater.

For outcomes of interest, manuscripts receive one point if they examine outcomes in four domains for which HomVEE historically has seen comparatively less research. Then for sample location, manuscripts

receive a half point if the entire study sample was in the United States at the time the study was conducted. Finally, for populations of interest, manuscripts receive a half point for each of the entire study sample being indigenous or the entire sample being one of MIECHV's priority populations named in the authorizing statute.

With the latest update, HomVEE adjusted the relative point value assigned to each manuscript for populations served. This was in response to stakeholder feedback during the public comment period that emphasized the importance of carefully considering which populations should get points from HomVEE at this stage. Specifically, we increased the points assigned to manuscripts that examine indigenous populations and MIECHV priority populations.

2. HomVEE Assigns Points to Each Model

After assigning manuscript-level points, the second step is to assign model-level points. These model-level points are based on criteria related to MIECHV program eligibility requirements. This increases the chance that models could be eligible for MIECHV funds, and that those models that could be eligible would be prioritized for review. We assign model-level points based on information from model websites, model developers, and information HomVEE has learned about the model in previous years of review.

Each model can earn up to four points. One point for being associated with a national organization or institution of higher education. These can be inside or outside of the United States. One point for being currently active or available to serve families. One point for being implemented for at least three years. Models can receive this point even if they are no longer currently active. Finally, one point if there's implementation support available in the United States. We assume that international models have implementation support in the United States if they've already been implemented in the United State or if the developers confirm that they would support implementation in the United States.

3. HomVEE Calculates Prioritization Scores for Each Model

After assigning manuscript model-level points, the third step is to calculate a prioritization score for each model. To do this, we first sum the manuscript-level points for all manuscripts about the model. At this stage, we group all manuscripts about related versions of a model together into one combined score. Models that have more eligible manuscripts then tend to receive more points. This prioritizes models with larger volumes of un-reviewed research. Next, we sum the manuscript-level point total and the model-level points. This becomes the unadjusted point total for the model. Then we adjust the model's point total, as I'll describe next. The model's final prioritization score is that point total multiplied by the adjustment factor.

Adjustment depends on the status and timing of the previous review

Now, the adjustment factor depends on the timing of the last review and which track the model is in. So, this slide describes three different weights that a model could receive. Models that were reviewed in the previous year receive a weight of zero, which lowers their prioritization score to zero and, therefore, a model will not be reviewed in two consecutive years. Models that were not reviewed in the prior year and are in track one, that is HomVEE hasn't identified those models as being evidence-based, those models receive a weight of one. So, these scores are equal to the unadjusted point totals based on manuscript and model-level points.

Finally, models not reviewed in the previous year that are in track two, those models that are evidence-based already, are adjusted by a weight based on the year they were last reviewed. This weight is higher for models that were reviewed less recently. That increases their scores relative to models that were reviewed more recently. In this way, HomVEE gives more priority to models that were reviewed longer ago when all other things are equal. So, for example, the prioritization score for a model considered for

review in 2021 that was last reviewed in 2017 would be multiplied by one plus one/tenth times 2021 minus 2017, all of that squared is 1.96. That's roughly double the score of the unadjusted point total.

4. HomVEE Gathers More Information on Top-Scoring Models and Updates Scores

After calculating the final adjusted prioritization score for all models, HomVEE sorts the models from highest to lowest separately in each track. Then the final step is to gather more information about top scoring models in each track and update those scores. For the top scoring models, we do three things. First, conduct a second focused database search looking only for model names to identify any additional manuscripts about the model that our search terms might not have captured. Any eligible manuscripts identified get added to the model's point total. Second, examining the full text of all of the eligible manuscripts about top scoring models to just confirm the top scoring model's points at the manuscript level. Finally, updating the model's prioritization score based on any newly identified research and the full text screening.

HomVEE uses prioritization points to select models to review

Last, using these final updated adjusted prioritization scores, HomVEE resorts the models separately in each track and then selects models for review, starting with the highest prioritization score and moving down the list in order. The number of models selected in each track depends on the project resources that HomVEE has available in any given year.

HomVEE Accepts Supplemental Information Only Under Specific Circumstances

Now, switching gears from prioritization, I wanted to cover a couple of procedural clarifications that might be of interest to HomVEE's users. One important procedural topic that HomVEE clarified in the version two handbook is the circumstances under which we accept supplemental information as part of the review. Supplemental information can take two forms, new information and new research. New information might be about a study's methods and procedures. HomVEE will incorporate this new information into the review if it's provided in direct response to an author query that HomVEE issues and is submitted in time.

If the author doesn't respond to an author query from HomVEE in time for the reviewers to examine that information during the same annual review cycle, then authors can wait until HomVEE releases the annual review results and, at that point, ask HomVEE to consider the new information through the reconsideration of evidence process. In contrast, new research could be additional findings or new analyses of research in a previously reviewed manuscript, or maybe could be an entirely new set of findings. HomVEE treats all new research as a submission to the next year's call for research unless it's requested by HomVEE in which case it would be incorporated into the review as I described earlier.

Subgroup Findings Are Only Reported If They Are Replicated

The second procedural clarification I wanted to go through relates to subgroup findings. Version two of the HomVEE handbook clarified the definition of subgroups and how subgroups would be reported. So, a subgroup is defined as a subset of the overall sample examined in the study. This is an analytic choice. It's different from defining a subgroup as a subset of the overall population. For example, if a study included mothers of all ages but also examined how the home visiting model affected teenage mothers specifically, HomVEE would treat the findings for teenage mothers as subgroup findings. However, if the overall study only enrolled teenage mothers, HomVEE would not treat the findings for teenage mothers as a subgroup finding.

Replicated subgroups. HomVEE defines as those with an identical definition in two non-overlapping research samples. So, for example, subgroups that are defined by cohort or time or location are not replicable, and HomVEE in the HHS criteria are really focused on replicated subgroups for a specific

reason. While researchers may examine an analytic subgroup because they're interested to learn about a subset of a population, the goal of the HHS criteria is to ensure that program impacts are replicated consistently for an outcome domain because this type of replication supports confidence that the evidence of effectiveness is not due simply to chance. So, HomVEE only reports subgroup findings from well-designed research if the subgroup is replicated. However, to help users understand what the research includes, HomVEE will provide a listing of the non-replicated subgroups that are included in well-designed research, beginning with the 2021 review results.

So, now I'll turn this over to Julieta to provide some more details about the specific standards that HomVEE uses to rate the quality of research that we review. That is how we assess our level of confidence that a home visiting model caused the observed impacts within research that HomVEE reviews.

Review of HomVEE Standards

Key Content

Thank you, Emily. Somebody mentioned the focus of this section is to provide you with an overview of the standards that have changed since version one. These changes include new validity and reliability standards, a revised baseline equivalence requirement, and changes in how we review the approaches with missing and imputed data in the methods of a statistical adjustment that are acceptable, and in how we review repeated measures, approaches, and structural equation models. As in the procedures section, a little blue badge is going to highlight the standards that have changed in version two.

Eligible Designs Include RCTs and QEDs

The eligible designs for review in HomVEE are randomized control trials, or RCTs, and quasi-experimental designs, or QEDs. These designs differ in how they assign members to different groups. So, that is RCTs assign them randomly, while in QEDs the assignment is not random. QEDs include single case designs. In these types of designs, the intervention and comparison conditions are assigned to a single family or a small group of families at different time periods.

In regression discontinuity designs, the effective intervention is estimated as the difference between mean outcomes of the intervention group members and comparison group members at a cutoff point. In non-experimental group designs, researchers created an intervention and a comparison group based on convenience or a statistical technique that matches sample members in each group. HomVEE used to refer to NEDs as match comparison group designs, but we adopted the new term, NED, to clarify that these types of studies can use a range of designs and are not limited to those that use the statistical matching techniques. In both RCT and NEDs, the impacts of a home visiting model are estimated by comparing outcomes between an intervention group and a comparison group. HomVEE does not consider simple comparisons of outcomes before and after the intervention to be eligible for review.

Eligible Comparisons

As we mentioned earlier, HomVEE rates findings from manuscripts, and these findings generally compare outcomes from an intervention group that received a home visiting intervention to outcomes for a comparison group that did not receive the intervention. The comparisons or contrasts that are eligible for reviewing HomVEE are those that allow us to answer HomVEE's core question, which is whether an early childhood home visiting model is effective. Because of this focus, a study then must examine the home visiting model in its entirety. If a design isolates a certain feature of a model, this design is not going to be eligible for review because knowing that a certain model feature is effective does not establish that a model consisted of these multiple features is effective overall.

More specifically, the comparisons that are eligible for review include comparisons in which the intervention condition is an early childhood home visiting model and the comparison condition is either any other intervention that is different to the home visiting model or the business as usual condition, which is the services routinely available to the population under study. With these types of designs, researchers are able to isolate the effect of the home visiting model, and that's why they are eligible.

The contrast that isolate the effect of a model feature or group of features are generally not eligible for review. This is, again, because knowing that a certain model feature is effective does not establish that the whole model is effective. Finally, contrasts that examine a combined home visiting models are not eligible either because they do not allow us to identify whether the effects are due to one of the models in particular or to the combination of the models.

Outcomes Must Now Demonstrate Face Validity and Reliability

Another new standard is that the outcomes measures must now demonstrate face validity and reliability. In version one, HomVEE only required that outcomes fall within one of HomVEE's eight outcome domains of interest. Face validity means that the measure is clearly described and that is appears to be measuring the concepts that is seeks to measure. Reliability indicates the extent to which the measure is actually measuring the concept accurately or consistently across different people or in different settings.

With this new standard, HomVEE ensures that findings are based on outcome measures that clearly and consistently measure concepts of interest to HomVEE. Findings that are based on measures that do not meet the face validity and reliability requirement will rate low. Reviewers may follow up with authors to query for this information and request citations.

Some Analytic Approaches Are Ineligible

Some analytic approaches are not eligible under the new standards. The approaches that are eligible are those that answer the main HomVEE question, which is whether a home visiting model is effective. Questions that focus on the mechanisms behind how a model works or the settings where a model might work best, and the populations who benefit the most from a particular model, which are important questions, will be outside of the scope of the HomVEE review and, therefore, not be eligible. This means that most mediating and moderating analyses which focus respectively on how or the circumstances under which a model works, or whether the model works equally well for different groups will not be eligible for review.

Additionally, all the analyses that control for endogenous characteristics will not be eligible. Endogenous characteristics describe participant behavior that emerges after a study participant learned their group assignment, or are behaviors that could be affected by a home visiting model. When the analysis used endogenous characteristics, we can no longer know whether an intervention or the change in behavior is creating the effect on outcomes. This creates bias estimates of intervention effectiveness.

So, as an example of what it means to include an endogenous variable, think of a study that measured the effects of home visiting services and children's language and literacy skills. If these analyses, at the same time that it's trying to estimate the effect of the home visiting model, is controlling for a variable, say, measure at halfway through the intervention that is trying to capture parent and child engagement, this engagement variable or characteristic is going to be picking up the effect of the home visiting model, because the home visiting model wants to influence parent-child engagement. So then the remaining -- we are going to be looking at an effect that is capturing both the effect of the intervention on the outcome but, at the same time, the effect of intervention of parent-child engagement, and the effect of parent-child engagement on the outcome. So, then our effect that we see of the intervention is biased because it's capturing these two effects at the same time. We don't know if it really the intervention or is this change in parent-child engagement that actually caused the effect in outcomes. Finally, HomVEE considers treatment-on-the-treated analysis to be eligible. However, intent to treat estimates are preferred and the HomVEE review will focus on the intent to treat estimate if available.

Revised Baseline Equivalence Requirements for RCTs and NEDs

The new standards also include that revised baseline equivalence requirement for RCTs and NEDs. We need to establish baseline equivalence so that we can make sure that if we observe an effect of a home visiting model, that effect can be attributed to the home visiting model and not to differences that existed between the intervention and the comparison groups. So, the new standards include the following. Low attrition RCTs no longer need to establish baseline equivalence or control for baseline variables to achieve a high rating. This is because RCTs use random assignment to create intervention and comparison groups. Because of that, the groups are expected to be equivalent across both measures and a baseline characteristic if attrition is low and randomization has not been compromised, all RCTs that have high attrition or have a compromised design.

NEDs still need to establish baseline equivalence to achieve a moderate rating. That is because whenever there is no randomization or the randomization is compromised, we cannot be sure that the two groups are equivalent on measure baseline characteristics, even if they are equivalent on measured baseline characteristics. So, the new standards still require that baseline equivalence of the analytic sample is established on baseline measures of the outcomes, on measures of race and ethnicity, and on socioeconomic status measures.

However, we have included two different changes. The first change is that maternal education was promoted to be a preferred measure of socioeconomic status. So, along with income, earnings, and poverty levels according to federal thresholds. The second change is that, to verify baseline equivalence, HomVEE will look now at effect sizes and will no longer look at a statistical significance of difference in means, that is it will no longer look at P values.

HomVEE Now Uses Effect Size to Verify Baseline Equivalence

An effect size measures the magnitude of the difference between intervention and comparison groups on a baseline characteristic, and it measures it in standard deviation unit. For example, an effect size equal to .14 will indicate that the mean value of a baseline variable for the intervention group is .14 standard deviations away from, which means below or above, the mean value of the baseline variable for the comparison group.

The way which HomVEE now looks at effect sizes is the following. If an effect size is greater than .25, that means then the intervention and comparison group are very different. So that means that this effect size is too big. So, the intervention and comparison group will not meet the baseline equivalence requirement and a finding that has this baseline difference will rate low. If a baseline effect size is equal or greater than .05, both lower or equal to .25, if the analyses includes an acceptable statistical adjustment for the baseline characteristic or outcome, then the finding can rate moderate. If a baseline effect size is lower than .05, then the finding will meet the baseline equivalence requirement and it will not require any statistical adjustment, and the finding can rate moderate. If the baseline data were imputed, then HomVEE will apply different standards that we're going to discuss in more detail later on.

Only Some Statistical Adjustments for Baseline Characteristics Are Acceptable

Only some statistical adjustments for baseline characteristics are acceptable. As noted in the previous slide, whenever the effect size is greater than .05 but less or equal to .25, studies can still meet the baseline equivalence requirement and then the findings can rate moderate if they use an acceptable statistical adjustment to account for these differences. The reason for this is that whenever the differences are too large to ignore, so that means in the .05 to .25 range, we can no longer be confident that the intervention, rather than these differences, are leading to the observed impact estimate. However, these differences are not so large that they cannot be accounted for within a statistical adjustment.

The acceptable analytic methods to adjust for baseline differences include regression adjustment, analysis of covariance or multivariate analysis of covariance. They also include estimating impacts only for groups defined at baseline. For example, ever had a baby compared to never had a baby. Repeated measures analysis of variance or multivariate analysis of variance. Please note that these approaches are subject to other requirements. Other acceptable approaches include growth curve modeling and if they meet additional conditions, such as having a correlation of .6 or higher and being the same or including the same measure of the outcome at baseline and follow-up, other methods are acceptable, which include gain or change of scores, difference-in-differences adjustment, and fixed effects for individuals.

Analyses with Missing Data Are Subject to New Standards

In version two standards, we have included a new approach to review analyses that include missing data. We have clarified which methods are acceptable for handling missing data. If analyses do not use one of these methods, then the findings from this analysis will rate low. The acceptable methods for handling missing data include complete case analysis, regression imputation, maximum likelihood estimation, and non-response weights. Using non-response weights only applies to missing outcome data but not to missing baseline data.

Similarly, it is acceptable to replace missing baseline data with a constant and combine these with including a missing data indicator. So, this method is not acceptable for missing outcome data, only for missing baseline data. Whenever HomVEE calculates overall and differential attrition rates, sample members with imputed outcome data will be counted as missing. In addition, version two of the WWC standards (Version 4.1) to review analytic approaches based on missing data.

Standards for Reviewing Analyses with Imputed Data

These new standards require that, in addition to using an acceptable method of imputation, QEDs and NEDs have to limit the potential bias from imputed outcomes data, and/or, if applicable, they must establish baseline equivalence using the largest baseline difference, accounting for missing or imputed data. The Version 2 HomVEE Handbook and the video link in this slide includes more details on the new approaches and the standards to review analysis with imputed data.

Repeated Measures Analysts Require Individual Time-Point Estimates

Another new standard applies to repeated measure analysis. In this type of analysis, researchers measure outcomes of the research sample at several points in time, with the goal of charting the sample growth over the course of the intervention and sometimes beyond the end of the intervention. In the new standard, HomVEE requires that findings from repeated measures analyses that include multiple follow-ups report findings for each of the individual time points. That is the analysis must report findings for each of the individual follow-ups. If these findings are not available, then the findings will not be considered eligible for review. Whenever the findings at each individual time point are available, HomVEE will assess and assign a separate rating to each finding at each individual time point. If the findings are not reported in the manuscript for each of the individual time points, then HomVEE will query authors to ask for this information at each time point. In case authors do not respond to the author query and the only available impact estimate is an impact estimate for combined time points, then this combined time point estimate will not be considered eligible for review.

Only Some Structural Equation Model (SEM) Analyses Are Eligible

Finally, another new standard in version two of the handbook applies to the structural equation models approaches. These are also known as SEM approaches. SEM is a statistical technique that examines the relationship between a dependent variable and two or more independent variables. To be eligible for

review, an SEM approach must present a path diagram for the model. If the path diagram for the model is not available, HomVEE will issue another query to request this path diagram.

In the path diagram, the researchers must include or indicate clearly that there is a direct pathway from the intervention to an outcome, and that there are no pathways leading to that outcome from another outcome. If this does not happen, then that particular estimate is not eligible for review. Another requirement is that the model must be identified. That is that the degrees of freedom must be greater than the number of parameters that are estimated in the model. If the model is not identified, then the findings are not eligible for review, either. Now I pass the presentation back to Shirley.

Rollout of Version 2 Procedures and Standards

HomVEE Will Apply Version 2 Procedures and Standards Beginning with the 2021 Review

Thank you so much, Julieta. I'll be discussing the rollout plans for version two procedures and standards. HomVEE plans to apply the version two procedures and standards beginning with the 2021 annual review. For models that are not evidence-based, HomVEE's plan is to review all research, and that includes any previously reviewed research, using version two at the time a model is prioritized for review using the process that Emily described earlier in the webinar.

For models that are already evidence-based, it's important to note that HomVEE will not generally retroactively apply the new standards to research that has already been reviewed by HomVEE. There are two exceptions to that. First, HomVEE will review previously reviewed, single case design research for any model selected for review using the new standards. This is because the requirements of the previous standards meant that single case design research reviewed to-date has never contributed to an evidence rating. This may now change.

Second, HomVEE will re-review subgroup research for any model selected for review in cases where there is research on a replicated subgroup but only a portion of that research was previously reviewed under the only standards. Because replicated subgroup findings must be considered together, if the subgroup replication straddled the timing of the update to HomVEE standards, HomVEE will review both sets of subgroup findings using the newest standards.

HomVEE will retroactively apply its clarified terminology, along with certain procedures, to all research. Specifically, in order to promote consistency in reporting across the review, clarifications about the outcomes and contrasts eligible for review in each domain, and the clarified definitions of study, manuscript, and subgroup will apply retroactively to all research on models, and this is true regardless of the model's evidence-based status, whether the model is prioritized for review, and whether HomVEE has reviewed the manuscript before.

Refer to The Handbook for Complete Details on The Revised Procedures and Standards

We do realize that this presentation has covered a lot of information. You can find detailed descriptions of HomVEE's revised procedures and standards in our handbook, which is published on the HomVEE website, and the link is presented on this slide.

Next Steps

I also wanted to mention that there are a couple of areas in which HomVEE is continuing to work on refining our procedures and standards, and we expect to publish updates on them in the future. First,

HomVEE has decided to continue refining our plans for defining and reviewing model versions. As we mentioned at the beginning of the webinar, the updates that are in the HomVEE Version 2 Handbook incorporated public comments based on two federal register notices we released in August. One of those notices focused specifically on the topic of proposed definitions and procedures related to model versions. This is a really important issue for HomVEE and for the field as a whole, and we really appreciate all of the very thoughtful comments we received. Based on those comments, we decided it would be best to take more time to refine our approach before moving forward with a new way of defining and reviewing model versions.

Second, HomVEE is continuing to work some updates related to the revised procedures and standards for single case design research. Some of the changes that we adopted in the version two handbook require updates to our procedures for applying the HHS criteria to single case design research. Currently, we're working with methodological and home visiting experts to develop an approach to this. Additionally, I want to mention that next month HomVEE expects to release a brief summarizing key public comments in response to the two federal register notices and HomVEE's approach to addressing them in the final version two handbook. We expect that to be available on our website in the coming month. Now we'll turn it over to Emily and we'd like to open up the Q&A.

Thanks, Shirley. Before I start addressing some of the questions, I'm just going to rewind a couple of slides, because I wanted to put back up the link to the handbook. For anyone who might be dialed in on audio and can't see the link, that's available at https://homvee.acf.hhs.gov/publications/methods-standards. I just want to make sure that folks can access that if they need to. The brief that Shirley mentioned will be available also in the publications area of our website when it's published. In addition, the recording of this webinar and its slides will be available on the publications area of the website, in the webinars section of the dropdown, in a couple of weeks.

Questions, Comments?

So, in terms of questions, we've received a few questions about our procedures and a few questions about the standards. So, we'll start with procedures questions. One question was, are MIECHV promising approaches also prioritized, and can promising approaches be reviewed in two or more years in a row? The answer to the first question is that the MIECHV program may coordinate with HomVEE to review promising approaches that are implemented and evaluated under a MIECHV grant. And the answer to the second question is that, just like other models, promising approaches cannot be reviewed two or more years in a row. I believe Shirley mentioned an exception to this, but I can't recall. In this first year that standards are rolled out, HomVEE is temporarily suspending that rule that a model has to wait for at least two years to be prioritized again after its been reviewed. So, models that were reviewed last year are eligible for prioritization this year when normally they wouldn't be.

A second question we received is, when a model is being reviewed, is there a way to know the cutoff for submitting items during the year that the model is being reviewed? So, for this, I'm just going to go back to an earlier slide about additional information. I'll share that one with you.

So, each year, HomVEE issues a call for research in around the middle of November, and it closes in early January. Anybody can submit any research during that time, and that's considered as part of the prioritization process in a following year's review. So, for example, in November 2020, we issued a call for research, and that closed at the beginning of January. All of that research will be considered as HomVEE prioritizes models for review in 2021. In addition, any research that came in from early January 2020 until November, when the call for research wasn't formally open but people may have sent us papers, that is also considered towards the 2021 review.

Then the exceptions for new information is if it's new information about a study's methods or procedures that HomVEE specifically asked for in an author query, that information will be incorporated into the current year review, not needing to wait a year. Otherwise, any new research is treated as a submission to the following year's call for study. For example, if HomVEE were to ask an author in March of 2021 to

provide some additional information about their study's methods and procedures, and the author wrote back and answered those questions but the author also attached an in press manuscript of a journal that's not available yet and said, "You may also be interested in this article," HomVEE would take that article and log it as a submission to the call for studies that closes in January 2021 and would start considering that as part of the 2022 prioritization process.

All right. So, now let's go to some of the questions about standards. So, one of the questions that we received about standards was to please explain the bullet on ineligible all analyses that control for endogenous characteristics again. That one's pretty complex, so, Julieta, I'll let you explain that, and I'll just go back a little bit to slide 35 where you were speaking about that point.

Yes, thank you, Emily. So, yes, what does it mean to have an endogenous variable or an endogenous characteristic? So, the way HomVEE understands endogenous characteristics is that these are behaviors, but these are behaviors that happen after the study participants learn about whether they are participating in the intervention or not. So, they learn about their treatment status. Or these are behaviors that can be affected by the home visiting model. So, if we include one of these endogenous variables as a control in, say, a regression that we use to estimate the impacts, we are controlling for something that is related to the intervention itself. So, then if we happen to see an impact of the intervention, we won't know for sure whether this impact occurred because of the actual intervention, which will be the home visiting model, or it occurred because the participants changed their behavior. So, that's why we do not want to include these variables.

So, in the example that I mentioned before, I can try to give more details and explain it better. So, think of an evaluation of a home visiting model that was implemented for two years. Think of this study as having a baseline assessment, an assessment at 12 months and an assessment at 24 months, that is after the baseline. Imagine that the main outcomes on these evaluations are to see the effects of the home visiting model on children's language and literacy skills at 24 months after the baseline. If, in this example, the researchers, in doing their analysis of the impacts of the home visiting model, also control for a variable that measures parent and child engagement, say in the middle of the implementation at 12 months, because you could think that parent-child engagement will be related to children's language and literacy skills, but if they happen to include that variable, parent and child engagement after 12 months of receiving home visiting services is going to be influenced by the home visit intervention. So then if we happen to serve an impact at 24 months after the baseline, then we really don't know whether that impact is due to the home visiting model or is it due because parent and child engagement improved either because of the home visiting services or because other factors that we don't know. So, when we see these, we consider the impact estimate to be biased because we cannot separate these effects. So that's why whenever we see a finding that comes from an analysis that includes an endogenous variable, HomVEE is not going to consider it eligible.

Thanks, Julieta. There is one other standards question here related to the baseline equivalence requirement. The question is, if the baseline equivalence is less than .05, why is it only rated moderate and not high?

So, I think this one is a little easier. So, what happens is that findings that rate high do not need to establish baseline equivalence. So, if they rate high, the assumption is that the differences in baseline measures are small enough. So, only findings that, at best, can get a moderate rating need to establish baseline equivalence. So, to get a rating of moderate, they need to meet the requirement. So, if they show that in their required baseline variables the effect size is lower than .05, then they meet the baseline requirement and they do not need to do a statistical adjustment for that baseline characteristic. But because they were only eligible for, at best, a moderate rating, then they will receiving the moderate rating.

Great. Thank you.

So, I hope that answers their question. I don't know if we have a way to provide additional information, if needed.

Well, we do have 15 minutes left in our webinar time and a couple more procedure questions came in while you were speaking. So, why don't I answer those, and if people have standards clarifying questions, they can send those in, or they might find the answers in the handbook afterwards. One of the questions that came in asks, if evidence-based models are changing their implementation practices, how does HomVEE take this into account when determining whether or not it still meets the criteria for its evidence base?

So, I think that this question might be about this slide. So, I'm going to answer what I think you're asking, but if this is not what you're asking, please write in to clarify. So, I think you're asking how we might assess whether implementation support is available if the model changes over time. And the answer to that is that we examine this again each time that a model has new research and we're considering it as a candidate for that year's prioritization. We look, again, at information on model websites, anything new the developer may have told us, and previous HomVEE reviews for this information.

One other simpler question that came in about procedures is when model information becomes available and posted on the HomVEE website. HomVEE releases the track one results each September, by September 30th, and that's for models that are not yet evidence-based, although, if they become evidence-based in that year, that September release is when that will be announced. Then the track two results for models that already had been evidence-based are released in December each year.

Let me go back over here and see what other questions may have come in. I think, Julieta, the question about baseline equivalence was answered. So, that's good news. I'm not seeing any additional questions here. I'll just pause for a moment and see if anybody wants to raise any questions. Okay. I'm not seeing any more coming in. If you have any further questions or comments, you can reach the HomVEE team at homvee@acf.hhs.gov. That contact information is also posted on the HomVEE website. Shirley, I just want to invite you to make any closing remarks that you might want to make as we end.

Thanks, Emily. I just want to take a moment to thank all our speakers today and all of our attendees for joining this webinar. We encourage you to visit the HomVEE website to read our handbook, review results from HomVEE, and learn more about the review. We really appreciate your time today and we hope you will visit us in the future. With that, this concludes our webinar for today.