CHILD AND FAMILY SERVICES REVIEW

TECHNICAL BULLETIN #13

Issued on August 6, 2021

This Technical Bulletin provides information on measurement and case review sampling for Round 4 of the Child and Family Services Reviews (CFSRs). This information is in accordance with, and provides guidance for, the following:

- Federal CFSR onsite review as required by 45 CFR § 1355.33(c);
- Criteria to determine substantial conformity with outcomes (45 CFR § 1355.34);
- Requirements that state Program Improvement Plans (PIPs) set forth the amount of progress [improvement] required, benchmarks used to measure state progress, and methods used to evaluate progress (45 CFR § 1355.35(a)); and
- The methods and information used to measure progress must be sufficient to determine when and whether the state is operating in subsequent substantial conformity or has reached the negotiated standard with respect to statewide data indicators that fail to meet the national standard for that indicator (45 CFR § 1355.35(e)).

Based on regulatory authority, we are making a limited number of adjustments to the statewide data indicators and case review sampling for CFSRs, and more substantial changes to measurement and sampling for PIPs. This technical bulletin has eight sections and three attachments. A brief description about each one is below:

- Section I describes how the CFSR statewide data indicators (SWDIs) will be used in the determination of substantial conformity for outcomes.
- Section II provides information regarding the case population and sampling methodology for CFSR onsite reviews.
- Section III provides information regarding measurement requirements for program improvement.
- Section IV provides information regarding the case review framework for PIP measurement.
- Section V describes the methodology that will be used to calculate the required amount of improvement and how we will evaluate state achievement of case review item measures.
- Section VI provides the methodology to calculate and determine achievement of measurement goals for the Item 1 aggregate data measure.
- Section VII describes the methodology that will be used to set PIP measurement goals and evaluate state achievement of required amount of improvement on the SWDIs.
- Section VIII provides information about the amount of time provided to assess progress and for states to achieve the required amount of improvement for case review item measures and SWDI included in the state's PIP Measurement Plan.
- Attachment A is the Round 4 data dictionary for the SWDIs and corresponding data quality checks. It provides indicator and data quality check descriptions and notations that identify Round 4 changes to two of the measures and several data quality checks.
- Attachment B provides technical information on the methodologies to calculate and categorize state risk-standardized performance (RSP) relative to national performance.
- Attachment C provides reference tables with examples showing required amounts of improvement for case review item measures.

Section I. Use of Statewide Data Indicators (SWDIs) to Determine Substantial Conformity on Outcomes

In CFSR Round 4, the determination of substantial conformity with Safety Outcome 1 and Permanency Outcome 1 will be informed by the state's risk-standardized performance (RSP) on the SWDIs in comparison to national performance *and* case practice ratings obtained through onsite case reviews.

The SWDIs are based on data submitted by states to the Adoption and Foster Care Analysis and Reporting System (AFCARS) and National Child Abuse and Neglect Data System (NCANDS). Indicator definitions for the SWDIs are provided in Attachment A, including notations to identify changes we are making for two indicators. Attachment B includes the technical steps to calculate and categorize states' RSP relative to national performance, which are generally the same as Round 3.

National performance is observed performance for the nation. We will calculate and publish national performance for each of the SWDIs prior to the start of CFSR Round 4, and those will remain fixed throughout the Round. A state's RSP will be compared to national performance for each SWDI to inform determinations of substantial conformity for the CFSR. Whether state performance is better, no different, or worse than national performance is based on a test of statistical significance. To make this assessment, we calculate 95% confidence interval estimates around each state's RSP for each indicator and compare the state's RSP interval estimate to the national performance. The state's RSP interval estimate must be better or no different than national performance on all applicable statewide data indicators associated with the outcome, and the state must meet the associated case review standards to be in substantial conformity with the outcome.

We will use the most recent RSP in the data profile that is transmitted to the state for the statewide assessment (approximately 6 months before the CFSR onsite review) to determine substantial conformity (i.e., RSP interval estimate is better or no different than national performance). We will update state performance data before making a final determination of substantial conformity if the state resubmits data to address data quality issues for applicable data periods by the specified resubmission due date.

Section II. Case Population and Sampling Methodology for CFSR Onsite Reviews

To ensure that the state child welfare case population (subject to review) meets federal guidelines, cases are sampled using a clearly defined sampling frame, based in part on the state's case review schedule and an adequate representation of the state child welfare system. In preparation for CFSR Round 4, we will work with states to develop and implement a state-specific case review sampling plan that reflects common requirements across all states.

A. Case Population

In accordance with 45 CFR § 1355.33(c), the CFSR requires an onsite review of a random sample of foster care and in-home services cases to evaluate the outcomes. The sampling frame is the list of cases from which random samples will be drawn. Similar to Round 3, the CFSR will consist of a minimum of 65 cases: 40 foster care and 25 in-home services cases. We are adjusting the in-home services case population subject to review for the CFSR and PIP in accordance with regulation at 45 CFR § 1355.35(c)(5).

The foster care sampling frame will consist of the listing of children served statewide or by jurisdiction strata according to the state's AFCARS-defined reportable case population. The foster care population for CFSR Round 4 will be the same as for Round 3.

The in-home services sampling frame will consist of non-foster care family cases opened 45 or more consecutive days to provide services and/or case management directly by the child welfare agency or through federally funded contract(s) (including alternative or differential response cases). The 45 days will begin the date the in-home services case was opened for case management and/or services. In Round 4, we will include family cases for children whose only placement setting during the sampling period was a trial home visit (THV) and the duration was 45 or more consecutive days from the start of the sampling period. Casework for children residing at home under the placement and care responsibility of the state is an important part of the child welfare continuum to assess. For CFSR Round 4, the in-home services case population will not include families who only received a Child Protective Services (CPS) investigation and/or assessment response. These changes will provide greater consistency across states in the in-home services case population and start date used to apply the 45-day case opening parameter.

B. CFSR Site Selection

CFSR case review sampling plans identify cases subject to review that reflect the state's child welfare population. The onsite review may be concentrated in several specific political subdivisions of the state, as agreed upon by the Children's Bureau (CB). One of the selected locations must include the state's largest metropolitan subdivision (45 CFR § 1355.33(c)(2)). The state's largest metropolitan subdivision is included as a site to ensure that the CFSR includes the urban center, where typically a larger proportion of families have contact with the child welfare system. In most states, the largest metropolitan subdivision is the entire county in which the state's largest city, by population, is located. State-specific information will be used when it is difficult to determine the state's largest metropolitan subdivision. The locality with the state's largest total child welfare case population (foster care and in-home services combined) should be a factor considered in site selection. The selection of sites is a collaborative process between states and the CB with the CB having final approval (45 CFR § 1355.33(c)(1) and (3)). See the *CFSR Procedures Manual* for more information regarding site selection for the CFSR onsite review.

C. Sampling Periods and Periods Under Review (PUR)

Similar to Round 3, 6-month sampling periods will be used for the foster care sampling frame, and 6 months plus 45 additional days for the in-home services sampling frame. In Round 4, the sampling period start date will begin 12 months prior to the start of the onsite review for both CFSR pathways (i.e., State-Led and CB-Led). The PUR starts at the beginning of the sampling period and ends when the cases are reviewed. In Round 4, sampling periods will only coincide with AFCARS submission periods (October 1 – March 31, and April 1 – September 30) for onsite reviews in the months of April and October. States conducting reviews across 3 or more months will be encouraged to use a rolling monthly or quarterly sampling approach (i.e., the sampling period advances one month/quarter per month/quarter of the review period) rather than a fixed approach to maintain a consistent PUR of 12 to 15 months. In Round 4, there will not be a requirement to review a balanced number of cases across sampling period intervals. See the *CFSR Procedures Manual* for more information regarding sampling periods and sampling frames for the CFSR onsite review.

Section III. Program Improvement Plan (PIP) Measurement Requirements

This section provides guidance on PIP measurement requirements that apply when a state is determined not to be in substantial conformity with Safety Outcomes 1 and 2, Permanency Outcome 1, and Well-Being Outcome 1. While states will not be expected to achieve a specified amount of improvement on

items rated as areas needing improvement for the other outcomes, we recommend states continue to assess performance across all domains. Child welfare practice in one domain impacts practice in the other domains and a comprehensive understanding of practice and system performance is critical to evaluating successful PIP implementation.

If we determine that a state is not in substantial conformity with an outcome due to its RSP on a SWDI, the state must include that indicator in its PIP and PIP Measurement Plan. We will consider the state's RSP performance for data periods after the statewide assessment and prior to PIP approval to determine final PIP and PIP Measurement Plan requirements. If we are unable to calculate a state's RSP on an indicator due to data quality issues, the state will be required to include that SWDI in its PIP and PIP Measurement Plan along with key strategies to correct the quality of the data. A summary of data quality checks and limits are included as Attachment C, including notations that identify changes for CFSR Round 4.

When a Safety Outcome is not in substantial conformity, the state's PIP Measurement Plan must include all of the safety SWDIs with RSP worse than national performance and all of the associated case review items rated as areas needing improvement as specified in the CFSR Final Report. This is consistent with guidance we have provided in prior rounds to comport with regulatory requirements for states to prioritize areas of nonconformity impacting child safety.

For Safety Outcome 1, there are two safety SWDIs and one case review item:

Statewide Data Indicators

- Maltreatment in Foster Care
- Recurrence of Maltreatment

Case Review Items

- Item 1: Timeliness of Initiating Investigations of Reports of Child Maltreatment

For Safety Outcome 2, there are two case review items:

Case Review Items

- Item 2: Services to Family to Protect Child(ren) in the Home and Prevent Removal or Reentry into Foster Care
- Item 3: Risk and Safety Assessment and Management

When Permanency Outcome 1 is not in substantial conformity, the state's PIP Measurement Plan must include all of the permanency SWDIs with RSPs worse than national performance as specified in the CFSR Final Report. All of the associated case review items rated as areas needing improvement for this outcome must be addressed in the PIP; however, there is no requirement that they be included in the state's PIP Measurement Plan. There are five SWDIs used to measure PIP performance for Permanency Outcome 1:

- Permanency in 12 months for children entering foster care
- Permanency in 12 months for children in care 12 to 23 months
- Permanency in 12 months for children in care 24 or more months
- Reentry to foster care
- Placement stability

When Well-Being Outcome 1 is not in substantial conformity, the state's PIP Measurement Plan must include each of the case review items for that outcome that were rated as areas needing improvement as specified in the CFSR Final Report:

- Item 12: Needs and Services of Child, Parents, and Foster Parents
- Item 13: Child and Family Involvement in Case Planning
- Item 14: Caseworker Visits With Child
- Item 15: Caseworker Visits With Parents

In situations where an updated data profile issued prior to PIP approval shows the state is either no different or better than national performance, the state will not be required to address that indicator in the PIP and PIP Measurement Plan.

Section IV. PIP Measurement and Sampling Framework for Case Reviews

This section describes the revised Round 4 framework for PIP case review measurement and sampling. We changed many components of the PIP measurement and sampling approach for Round 4 to strengthen regulatory intent to evaluate the PIP and measure a state's progress toward substantial conformity (§ 1355.35(a) and (e)), and to integrate measurement principles for using qualitative case reviews to detect change in performance.

A. Case Population

The case population subject to review for PIP measurement is the same population reviewed for the CFSR; see Section IIA of this TB.

A Round 4 change to the case elimination criteria is to eliminate cases that were reviewed in the previous 12 months. This change is to reduce potential burden placed on families and key participants to be involved in case participant interviews more than once a year.

B. PIP Measurement Sites

For purposes of PIP measurement, states will identify a select group of sites or localities designated in the PIP as target populations for implementation of selected strategies and interventions. These localities are often referred to as PIP implementation sites. Selection of PIP measurement sites is a collaborative process between states and the CB, with the CB having final approval (§ 1355.35(a)(2)). Selection of implementation sites will be guided by evidence used to develop the PIP (including evidence provided in the statewide assessment and the CFSR Final Report), selected PIP strategies and interventions, demonstrated practice improvement needs, readiness to be an implementation site, and the CQI change and implementation process. States may propose all or a subset of PIP implementation sites to be PIP measurement sites. Selecting a group of implementation sites targeted in a state's PIP to be PIP measurement sites will provide greater confidence that change detected during PIP measurement is likely the result of PIP implementation strategies and interventions.

C. Baseline Period

The amount of improvement a state will be required to attain on each case review item in the PIP Measurement Plan will be calculated based on performance during the baseline period. Baselines will be established post-CFSR, after PIP implementation sites are selected and approved by the CB. The baseline period will be measured as Measurement Period 1 to encourage continuous measurement during PIP implementation and to help reduce the amount of time between the baseline period and subsequent

PIP measurement periods. These changes better align PIP measurement with principles for using qualitative case reviews to detect change and provide greater confidence that changes detected are the result of PIP implementation. We recommend that the state complete Measurement Period 1 (baseline) within the first 6 months of the PIP Implementation Period, and it must be completed within 12 months. CB may approve a state beginning PIP measurement post-CFSR before the PIP is officially approved by CB if the state is prepared, PIP measurement sites have been identified and approved by CB, and the PIP Measurement Plan has been endorsed by the CB Measurement and Sampling Committee (MASC) and approved by the CB Regional Office.

D. Measurement Periods

For case review items included in a state's PIP Measurement Plan, we will evaluate attainment of the required amount of improvement based on performance after Measurement Period 1 (i.e., post-baseline). In CFSR Round 4, each measurement period will consist of a unique population of cases. We will discontinue the process of advancing measurement periods that include some cases from the prior measurement period. We recommend states use measurement periods of 6 months or less to maximize the number of opportunities a state has to achieve the required amount of improvement for PIP measurement, assess progress, and make needed adjustments. Shorter measurement periods will also provide repeated sampling results, which are recommended when using small samples to detect change in performance. The maximum length of a measurement period will be 12 months. States opting to use a 12-month measurement period may be limited to two measurement periods to achieve the required amount of improvement unless the state is approved to begin measurement before the PIP is officially approved and/or a condensed schedule is used for the final measurement period.

E. Sample Size and Minimum Applicable Cases

States electing to use recommended measurement periods of 6 months or less will need to review a sample of cases that result in a minimum of 33 applicable cases (20 foster care and 13 in-home services cases) per case review item each measurement period. States electing to use measurement periods of 7 to 12 months will need to review a minimum of 65 cases (40 foster care and 25 in-home services cases) with at least 33 applicable cases per case review item.

Requiring at least 33 applicable cases be reviewed per item provides a reasonable sample size to calculate the required amount of improvement and draw inferences from repeated measurement about practice. Similar to Round 3, the number of applicable cases reviewed for each item during Measurement Period 1 (i.e., the baseline) will be the minimum number of applicable cases required for each subsequent measurement period to calculate performance. We will continue to provide a 2% tolerance when determining the minimum number of applicable cases required for ongoing measurement. This is calculated by multiplying the number of applicable cases for the item during Measurement Period 1 by 0.02 and rounding to the nearest whole number.[1] The resulting value is subtracted from the number of applicable cases reviewed for that item during Measurement Period 1 to identify the minimum number of applicable case reviews required to calculate performance in subsequent measurement periods. The 2% tolerance will apply even if the resulting minimum number of applicable cases is below the minimum 33 cases.

[1] If the number behind the decimal point is less than 5, the value is rounded down to the nearest whole number. If the number behind the decimal point is equal to or greater than 5, the value is rounded up to the nearest whole number.

F. Stratification

We recommend that PIP measurement samples reflect a similar distribution of the child welfare case population as the PIP measurement sites, when possible. We encourage states to limit the number of factors used to stratify the sample. States should consider the proportion of cases in the locality with the largest child welfare case population, often the metropolitan area, in comparison to the proportion of cases in the state and other PIP measurement sites. Similarly, we will negotiate the ratio of foster care to in-home services case types by considering the sample size and case type proportions in the state and PIP measurement sites. Proportions used for the baseline period will need to be similar in subsequent measurement periods. Similar to Round 3, we will provide a +/-5% tolerance to the proportion of cases reviewed in the metro site or site with the largest case population and by case type when comparing the baseline period with subsequent measurement periods. The tolerance is applied by adding and/or subtracting 5 percentage points from site and case type proportions established during Measurement Period 1. State proposals to use additional strata will be negotiated individually, based on the number of additional strata proposed, rationale, and the case review sample size.

G. Sampling Periods and Periods Under Review for PIP Measurement

Similar to the CFSR, sampling periods for PIP measurement will be 6 months for foster care sampling frames and 6 months plus 45 days for in-home services sampling frames. States will be encouraged to use a rolling monthly or quarterly sampling approach (i.e., the sampling period advances one month/quarter per month/quarter of the review period) to maintain a consistent PUR. For PIP measurement in Round 4, we recommend states use a 7- to 9-month PUR. Shorter PURs reflect more recent case practice and are more likely to detect changes from implementation of PIP strategies and interventions. There will not be a requirement to review a balanced number of cases across sampling period intervals. We will consider the use of a fixed sampling approach when requested.

Section V. Methods to Calculate and Determine States Achieved Required Amount of Improvement for Case Review Items

This section describes the revised Round 4 PIP measurement methodology to calculate and determine if states achieved the required amount of improvement for case review items. The required amount of improvement will be determined based on the state's performance on each case review item during Measurement Period 1 (i.e., the baseline). States will have two pathways to achieve the required amount of improvement. The first pathway is to meet or exceed the sustained improvement goal in any three measurement periods. This pathway is new in Round 4 and reflects regulatory intent to evaluate the PIP and measure a state's progress on their PIP (§ 1355.35 (a)(1)(v) and (e)(1-2)) by encouraging sustained performance across multiple measurement periods. It also aligns the use of qualitative case reviews for PIP measurement with qualitative research principles of data saturation and reliance on multiple data points to detect change using small samples. The second pathway is to meet or exceed the high-performance value in any single measurement period. This pathway provides states the ability to achieve the required amount of improvement using a single measurement period. The high-performance value is set well above the sustained improvement goal to strengthen confidence that change in performance is detected using a single measurement point. Achievement of the required amount of improvement for case review items will be determined by whichever pathway is achieved first.

A. Method to Calculate Required Amount of Improvement

This section describes how we will calculate the required amount of improvement for each case review item requiring inclusion in the state's PIP Measurement Plan. The required amount of improvement is expressed as a percentage and will be adjusted based on a state's baseline performance on that case review item. The higher a state performs during Measurement Period 1 (i.e., the baseline), the smaller the amount of improvement required; the lower a state performs, the larger the amount of improvement required. The amount of improvement is scaled to provide a reasonable amount of improvement that differentiates across the range of baseline performance amounts. The formula to calculate the required amount of improvement is:

=100(0.25(Percentage Equal to a CFSR Strength Rating – Baseline Performance))

Baseline performance for an individual item is calculated as:

=Number of Strength Ratings / Number of Applicable Cases

Step 1. Calculate a state's performance at baseline.

A state's baseline performance is calculated by dividing the number of Strength ratings for that item by the total number of cases during Measurement Period 1 that are applicable to that item. To determine if ongoing measurement is required, the performance is rounded to the nearest whole number. If the state's performance is equal to or exceeds the percentage of applicable cases that would equal a CFSR Strength rating, we will consider the required amount of improvement attained and ongoing measurement will not be required; for Item 1, this is 95%; for all other case review items, this is 90%. The baseline performance is not rounded when used to compute the required amount of improvement.

Step 2. Calculate the required amount of improvement.

Calculate the scaling factor by subtracting the state's unrounded baseline performance from the percentage of applicable cases that would equal a CFSR Strength rating. For Item 1, this is 0.95 (95%). For all other case review items, this is 0.90 (90%). Multiply the scaling factor by 0.25 to adjust the scale and calculate the required amount of improvement. Multiple this value by 100 to express the amount of improvement as a percentage.

In Round 4, the required amount of improvement is scaled and determined by the state's baseline performance. By subtracting a state's baseline from the percentage of applicable cases that would equal a CFSR Strength rating, the higher a state performs at baseline, the smaller the resulting scaling factor; the lower a state performs at baseline, the larger the scaling factor. The scale is adjusted by multiplying the scaling factor by 0.25. This amount represents a scale of 23.75 percentage points (i.e., 0.25 × 95% for Item 1) or 22.5 percentage points (i.e., 0.25 × 90% for all other case review items) distributed across all mathematically possible baseline performance values. The range of possible baseline performance values is dependent on the number of applicable cases. Using 0.25 results in a reasonable distribution in the required amount of improvement that differentiates across the range of possible baseline performances. For example, for 65 applicable cases, approximately every 4 additional Strength ratings at baseline will result in one less additional Strength rating needed to achieve the required amount of improvement during subsequent measurement periods.

Step 3. Calculate the sustained improvement goal.

The sustained improvement goal for each item is calculated by adding the amount of improvement calculated in Step 2 to the state's unrounded baseline performance and is expressed as a percentage. The sustained improvement goal is rounded up to the nearest whole number.[2] The sustained improvement goal is based on each state's baseline performance and will be capped at the percentage of applicable cases that would equal a CFSR Strength rating. For Item 1, this is 95%; for all other case review items, this is 90%.

Step 4. Calculate the high-performance value.

The high-performance value is calculated similarly to the sustained improvement goal, using the formula:

=100(0.50(Percentage Equal to a CFSR Strength Rating – Baseline Performance))

Repeat Step 1, and in Step 2, multiply the scaling factor by 0.50 and add that amount to the state's unrounded baseline performance to calculate the high-performance value. This value is also expressed as a percentage and rounded up to the nearest whole number. Similar to the sustained improvement goal, the high-performance value will be capped at the percentage of applicable cases that would equal a CFSR rating. For Item 1, this is 95%; for all other case review items, this is 90%.

Appendix D displays examples of sustained improvement goals and high-performance values calculated across various baseline performances for different numbers of applicable cases.

B. Method to Determine States Achieved Required Amount of Improvement

This section describes how we will determine a state attained the required amount of improvement for each case review item requiring measurement in a state's PIP Measurement Plan. States report achievement of the required amount of improvement in PIP Progress Reports, which we will confirm using a verification process. We will calculate state performance by dividing the number of Strength ratings for an item by the total number of case reviews applicable to that item for the measurement period. Performance will be rounded to the nearest whole number. To evaluate the attainment of the required amount of improvement, measurement periods must meet all applicable measurement criteria, including the minimum number of applicable cases and distribution of cases by case type and site with the largest case population.

States will have two pathways to meet the required amount of improvement for each case review item: (1) meet or exceed the sustained improvement goal in any three measurement periods, *or* (2) meet or exceed the high-performance value in any single measurement period. States will not be required to select a pathway. Achievement of the required amount of improvement for case review items will be determined by whichever pathway is achieved first.

Pathway 1. Meet or exceed the sustained improvement goal in any three measurement periods.

We will determine the measurement requirement for a case review item attained when a state reports performance for three measurement periods that meet or exceed the sustained improvement goal for the item. There is no requirement that the three measurement periods be consecutive. Once achieved,

[2] If the numbers behind the decimal point are greater than 0, the value is rounded up to the nearest whole number.

ongoing PIP measurement of that case review item will not be required. The requirement for repeated performance measurement at or above the sustained improvement goal increases confidence that change in performance can be detected using small samples for qualitative case reviews. States that elect to use 12-month measurement periods may not have a sufficient number of measurement periods to achieve the required amount of improvement in this manner.

Pathway 2. Meet or exceed the high-performance value in any single measurement period.

We will determine the measurement requirement for a case review item attained when a state reports performance that meets or exceeds the high-performance value in any single measurement period. Once the high-performance value is achieved, ongoing PIP measurement of that case review item will not be required.

Section VI. Method to Calculate and Determine Achievement of Required Amount of Improvement for Item 1 State Aggregate Data Measure

We recommend that states required to measure Item 1, Timeliness of Initiating Investigations of Reports of Child Maltreatment, in their PIP Measurement Plan use a statewide aggregate data measure. This section describes how we will calculate and determine the required amount of improvement for states electing this option. For the Item 1 Aggregate Measure, the denominator is defined as the number of children identified in accepted/screened-in CPS reports that require face-to-face contact in a 12-month period. The numerator is defined as the number of children identified in the denominator that received timely face-to-face contact according to agency policy. Face-to-face contact is calculated as the difference between two dates: the date the CPS report is accepted/screened-in and the date of face-to-face contact with the child. This amount of time is compared to the state's required timeframe(s) for each report's priority level to determine whether the contact was completed timely. The resulting proportion is expressed as a percentage. The baseline period and subsequent measurement periods will comprise 12 consecutive months of practice. The baseline period may begin on or after the sampling period start date used for the CFSR onsite review.

A. Method to Calculate Required Amount of Improvement

This section describes how we will calculate the required amount of improvement for the Item 1 Aggregate Measure. The required amount of improvement is expressed as a percentage and will be adjusted based on a state's baseline performance. The maximum possible required amount of improvement will be 3.0% above the state's baseline performance. A maximum improvement amount of 3.0% was determined to be a reasonable amount of improvement for this aggregate measure and decreases the higher a state performs above a baseline performance of 50.0%. 50.0% was selected as the starting point for scaling as no state that elected to use a statewide aggregate data measure in Round 3 had a baseline performance below 50.0%. Because the Item 1 Aggregate Measure uses the entire population of children identified in accepted/screened-in CPS reports that require face-to-face contact, there are no sampling concerns. The formula to calculate the required amount of improvement is:

$$=100(0.03((1.0 - \text{Baseline Performance}) + 0.5))$$

Baseline performance for the Item 1 Aggregate Measure is calculated as:

=Number of Timely Face-to-Face Child Contacts / Number of Face-to-Face Child Contacts Required

Step 1. Calculate a state's baseline performance.

A state's baseline performance is calculated by dividing the number of children receiving timely face-to-face contacts, in accordance with state policy regarding priority level timeframes, by the total number of children requiring face-to-face contacts for the 12-month period. To determine if ongoing measurement is required, baseline performance is rounded to the nearest tenth of a percent. If the state's baseline performance is equal to or exceeds the percentage of applicable cases that would equal a CFSR Strength rating (95%), we will consider the required amount of improvement attained and ongoing measurement will not be required. The baseline performance is not rounded when used to compute the required amount of improvement.

Step 2. Calculate the required amount of improvement.

To determine the required amount of improvement, we calculate a scaling factor by subtracting the state's unrounded baseline performance from 1.0 and adding 0.5. The maximum scaling factor value is capped at exactly 1. Multiply the scaling factor by 0.03 to calculate the required amount of improvement. Multiply this value by 100 to express the amount of improvement as a percentage.

By subtracting a state's baseline from 1.0, the higher a state performs at baseline, the smaller the resulting scaling factor; the lower a state performs at baseline, the larger the scaling factor. Adding 0.5 adjusts the scale to account for a minimum expected baseline performance of 50.0%. The maximum value of the scaling factor is capped at exactly 1 to subsequently cap the maximum required amount of improvement at 3.0%. Based on available Round 3 data, we do not expect a state's baseline performance below 50.0%. Setting a cap on the maximum amount of improvement prevents an unrealistic amount of improvement for any potential baseline performance less than 50.0%.

Step 3. Calculate the item measurement goal.

Add the amount of improvement calculated in Step 2 to the state's unrounded baseline performance to calculate the item measurement goal. The item measurement goal is rounded to the nearest tenth of a percent. The goal is capped at 95%, which is the percentage of applicable cases that would equal a CFSR Strength rating. Table 1 displays the scaling factor, amount of improvement, and item measurement goal for various baseline performances.

Table 1. Scaling Factors, Required Amounts of Improvement, and Item Measurement Goals for a Range of Baseline Performances for the Item 1 Aggregate Measure

Baseline Performance	Scaling Factor	Required Amount of Improvement	Item Measurement Goal
0.0%	1	3.0%	3.0%
5.0%	1	3.0%	8.0%
10.0%	1	3.0%	13.0%
15.0%	1	3.0%	18.0%
20.0%	1	3.0%	23.0%
25.0%	1	3.0%	28.0%
30.0%	1	3.0%	33.0%
35.0%	1	3.0%	38.0%
40.0%	1	3.0%	43.0%

Baseline Performance	Scaling Factor	Required Amount of Improvement	Item Measurement Goal
45.0%	1	3.0%	48.0%
50.0%	1	3.0%	53.0%
55.0%	0.95	2.9%	57.9%
60.0%	0.90	2.7%	62.7%
65.0%	0.85	2.6%	67.6%
70.0%	0.80	2.4%	72.4%
75.0%	0.75	2.3%	77.3%
80.0%	0.70	2.1%	82.1%
85.0%	0.65	2.0%	87.0%
90.0%	0.60	1.8%	91.8%
95.0%	N/A	N/A	N/A

Note: N/A indicates not applicable as ongoing measurement is not required.

B. Method to Determine State Achieved Required Amount of Improvement

Following the 12-month baseline period, state progress will be measured using distinct, consecutive 12-month measurement periods. For example, if the 12-month baseline is set using the period January 2022 to December 2022, the next measurement period for the Item 1 Aggregate Measure will be January 2023 to December 2023, followed by January 2024 to December 2024. We will determine the required amount of improvement for the Item 1 Aggregate Measure attained when a state meets or exceeds the item measurement goal in any single measurement period. Once achieved, ongoing measurement of Item 1 will not be required.

We will apply a high-performance plateau adjustment when an aggregate measure is used for Item 1 and the PIP measurement goal is set at 90% or above. In these situations, if the state is able to sustain performance at or above the baseline for three measurement periods, we will consider the goal met, even if the state does not meet the actual goal.

Section VII. Methods to Establish PIP Measurement Goals and Evaluate Required Amount of Improvement Achieved for Statewide Data Indicators

We will implement the same methods intended for Round 3 to identify baselines and calculate improvement goals for statewide data indicators that require PIP measurement. The data periods used to establish baselines and goals will vary by state based on the CFSR schedule. States will have two pathways to achieve the required amount of improvement for each SWDI included in the PIP Measurement Plan. The first pathway is for the state's observed performance for a 12-month reporting period to meet or exceed the improvement goal. The second pathway is for the state's RSP for a 12-month reporting period to be better or no different than national performance. Achievement of the required amount of improvement for SWDIs will be determined by whichever pathway is achieved first.

A. Baselines

Baselines will be established using the most recent observed performance in the data profile available at the time the CFSR Final Report is completed. Observed performance is used rather than RSP because PIP measurement evaluates state performance over time relative to that state and not relative to the nation. The most recent observed performance is used as the baseline rather than an average of past

performance because that could result in a goal that is worse than state performance, for instance, if a state made a considerable amount of improvement in the most recent reporting period. Conversely, using an average of past performance for states that experienced a decline in performance in the most recent reporting period could result in an improvement goal that is difficult for the state to achieve. Baselines will remain fixed throughout PIP measurement except for situations when a state is approved to resubmit data for the baseline period.

B. Method to Calculate Statewide Data Indicator Improvement Goals

Improvement goals for the SWDIs are calculated using an improvement factor that is based on the state's past 3 years of observed performance. Three years of data are used to account for variability in state performance. The improvement factor is multiplied by the state's observed performance for the most recent year of observed performance at the time the CFSR Final Report is completed (baseline). The improvement factor may be decreased or increased based on national floors and caps that are set and fixed throughout Round 4 for each indicator. Caps and floors are used to address improvement goals that would require either a large amount of improvement that could be difficult to reach, or too little of amount of improvement to be meaningful. The cap is set so that no state is required to improve by more than the amount of improvement equal to the value of the improvement factor at the 50th percentile. Floors are set so states are required to improve by at least the amount of improvement at the 20th percentile or 80th percentile, depending on whether higher or lower performance is desired. States with an improvement factor that exceeds the cap will have the factor replaced with the cap. States with an improvement factor below the floor will have the factor replaced with the floor. Improvement factors that fall between the cap and floor will use the improvement factor generated by the state's data to calculate the improvement goal.

The method to establish improvement goals is described in the following steps:

Step 1. Calculate 7 estimates of past performance.

We use a state's data from the past 3 years (Years 1, 2, and 3) to obtain 7 values: the state's actual performance in Year 1, Year 2, and Year 3, and the averages of Years 1 through 3, Years 1 and 2, Years 1 and 3, and Years 2 and 3. Using all seven values instead of the original three provides a more reliable estimate of the state's overall past performance and amount of fluctuation that is typical for that state. As an example, consider that in the most recent three 12-month reporting periods, State Y discharged 36.6%, 41.6%, and 36.8% of the children who entered foster care to permanency within 12 months. From this we create 7 estimates as displayed in Table 2.

Table 2. Example showing calculation of 7 estimates of past performance

Year 1	Year 2	Year 3	Years 1–3 (Average)	Years 1 & 2 (Average)	Years 1 & 3 (Average)	Years 2 & 3 (Average)
36.6%	41.6%	36.8%	38.3%	39.1%	36.7%	39.2%

Step 2. Estimate the variability in past performance.

Some states, particularly larger states, show small fluctuations in performance from year to year. Other states show larger changes in performance over time, often due to smaller and more variable population sizes. To account for the state's variability in past performance, the standard deviation (SD) is used. In our case, the SD will tell us how much our estimates of past performance vary from the average

(mean) of these estimates. A low SD indicates very little variability—the data points tend to be very close to the mean; a high SD indicates a lot of variability—the data points tend to be spread out over a larger range. To create more reliable and accurate estimates used to account for the underlying variability in each state's past performance, we take the following steps:

Step 2.1. Calculate a sample mean and standard deviation, based on the 7 values of past performance.

The mean reflects an estimate of each state's past performance based on the past 3 years of observed performance, and the SD reflects the overall variation in this past performance. The SD tells us how far away the 7 values tend to be from the mean. To increase the reliability and accuracy of the estimates, a technique called bootstrapping is used. Bootstrapping involves taking the original 7 values and sampling with replacement to generate a large number of resamples of the 7 values. The process does not create any new numbers: each resample is simply a random sample selected from those same 7 values. Because we resampled with replacement, some of the 7 values are repeated in each of the 1,000 samples and there is no guarantee that any of the 1,000 resamples will contain all of the 7 values.

Step 2.2. Bootstrap the mean and standard deviation 1,000 times.

We repeat the bootstrap process 1,000 times to create 1,000 sample means and 1,000 sample standard deviations. Repeating the process many times is a common technique of bootstrapping, and 1,000 is a reasonable number of bootstraps. This step does not increase the amount of information in the original dataset as it is done using the 7 values calculated in Step 1.

Step 3. Calculate a grand mean and mean standard deviation.

The mean of the 1,000 sample means is calculated to generate a single value and is referred to as the grand mean, and the grand mean represents the state's past performance on the indicator. The mean of the 1,000 SDs is also calculated to generate a single value that represents the state's typical variability in past performance for the indicator and is referred to as the mean standard deviation (MSD).

Table 3. Example showing grand mean and mean standard deviation

Grand Mean	Mean Standard Deviation (MSD)
38.3%	1.8%

Step 4. Calculate the improvement factor.

An improvement factor is calculated by adding two mean standard deviations (MSD) to the grand mean and dividing that amount by the grand mean. The improvement factor is rounded to 3 decimal places.

Using the example with a grand mean of 38.3% and MSD of 1.8% obtained by bootstrapping, we calculate an improvement factor of 1.094—see Table 4 below.

Table 4. Example showing calculation of improvement factor

Year 1	Year 2	Year 3 (Baseline)	Grand Mean	Mean Standard Deviation (MSD)	Grand Mean + 2 MSDs	Improvement Factor
36.6%	41.6%	36.8%	38.3%	1.8%	41.9% (38.3% + 2*1.8%)	1.094 (41.9% / 38.3%)

As mentioned in part A of this section, the most recent year of observed performance is used as the baseline, rather than the grand mean, to avoid setting goals that are either out of reach or are worse than the most recent year. The improvement factor is calculated using the grand mean and MSD, but it is applied to the observed performance for the most recent year (baseline). Therefore, the improvement factor is applied to the baseline as a multiplier (baseline * improvement factor = improvement goal). Using the example in Table 4, this would result in an improvement goal of:

36.8% * 1.094 = 40.3%

Step 5. Compare the state's improvement factor to the national cap and floor for improvement factors to determine whether to use an adjusted improvement factor (cap or floor).

Occasionally, using an improvement factor will yield performance goals that would require a large amount of improvement that could be difficult to reach, or conversely, too small an amount of improvement to be meaningful. For instance, in the example above, the improvement goal of 40.3% requires a nearly 10% increase over the baseline. To address this issue, we will set a national cap and a floor for improvement factors for each indicator and use those when a state's improvement factor exceeds the national cap or floor. The national cap and floor are generated using the distribution of all states' improvement factors, and the national cap and floor for each indicator will be set prior to and remain fixed throughout Round 4.

The method to establish the national cap and floor for each indicator is identified in the steps below:

Step A. Calculate improvement factors for all states.

Calculate improvement factors for *all* states (regardless of whether the state is required to include the SWDI in the PIP), and rank and re-order the states according to the size of the improvement factor, from largest to smallest.

Step B. Determine the national cap for improvement factors (50th percentile).

For each indicator, determine the middle of the distribution (50th percentile) of the set of improvement factors for all states. This will be the cap on the improvement factor. If a higher percentage is the desired direction of performance (as in the example here), improvement factors greater than the one at the 50th percentile will be replaced with the one at the 50th percentile. Conversely, if a lower percentage is the desired direction of performance, improvement factors lower than the one at the 50th percentile will be replaced with the one at the 50th percentile. This will be referred to as an "adjusted improvement factor."

Step C. Determine the national floor for improvement factors (20th or 80th percentile).

Determine the 20th percentile for indicators where higher performance is desired (i.e., permanency in 12 months indicators) and determine the 80th percentile for indicators where lower performance is desired (i.e., reentry to foster care, recurrence of maltreatment, maltreatment in foster care). This will be the floor on improvement factors. The purpose of establishing a floor at the 20th or 80th percentile is to set a minimum limit on the required amount of improvement.

Step 5.1. For states with an improvement factor beyond the floor, replace it with the value at the 20th or 80th percentile (depending on the indicator) to obtain the "adjusted improvement factor."

Step 5.2. For states with an improvement factor beyond the cap, replace it with the value at the 50th percentile to obtain the "adjusted improvement factor."

Step 5.3. For states with an improvement factor that falls between the cap and floor, use the original improvement factor generated from the state's data.

Step 6. Calculate the improvement goal.

Step 6.1. Round the baseline to one decimal place if it is an indicator expressed as a percent and round it to two decimal places if the indicator is expressed as a rate. Round the improvement factor from Step 5 to three decimal places.

Step 6.2. Multiply the rounded improvement factor (as applicable) to the rounded Year 3 observed performance value (baseline) to calculate the improvement goal.

Table 5. Example showing calculation of improvement goal

Year 1	Year 2	Year 3 (Baseline)	Improvement Factor	Example *Adjusted* Improvement Factor	Improvement Goal
36.6%	41.6%	36.8%	1.094 (41.9% / 38.3%)	1.063 (*example* cap) *To be determined based on national calculations for Round 4*	39.1% (36.8% * 1.063)

C. Method to Determine State Achieved Required Amount of Improvement

A state can achieve the required amount of improvement on the SWDIs for PIP measurement in one of two ways. Achievement of the required amount of improvement will be determined by whichever pathway is achieved first.

Pathway 1. Observed performance meets or exceeds the improvement goal.

Pathway 2. Risk-standardized performance is better or no different than national performance.

We will evaluate state performance on the indicators using the state's data profile that is published approximately every 6 months. Performance on the permanency indicators will be assessed every 6 months using the state's available AFCARS data. Performance on the safety indicators will be assessed every 12 months using the state's available NCANDS data, and AFCARS data as applicable.

Section VIII. Amount of Time Provided for PIP Measurement

This section pertains to the evaluation of state progress and attainment of the required amount of improvement for PIP measurement.

Similar to Round 3, states will be provided the 2-year PIP Implementation Period followed by an 18-month Post-PIP Evaluation Period (previously referred to as the non-overlapping evaluation period) to attain the required amount of improvement on measures included in the PIP Measurement Plan.

While all states will have no less than 42 months to meet PIP measurement requirements, some states may have additional time to meet the required amount of improvement for case review items if they obtain a MASC-endorsed and CB Regional Office-approved PIP Measurement Plan, including approval for case review PIP measurement sites, and begin Measurement Period 1 before the start of the PIP Implementation Period.

During the 18-month Post-PIP Evaluation Period, some states will receive one data profile with new calculations for state performance on the two safety indicators, while other states will receive two. Performance on the safety indicators requires NCANDS data, which are submitted by states annually.

Therefore, published data profiles have updated performance values for safety indicators annually, in August, after the newly submitted NCANDS data are received and validated. We will calculate and transmit updated safety indicator performance data within 6 months following the end of the Post-PIP Evaluation Period for the states that receive only one profile with new safety data indicator performance during the Post-PIP Evaluation Period. We will consider those states' safety indicator performance data for the most recent reporting period when making a final determination regarding whether the required amount of improvement specified in the PIP Measurement Plan was achieved.

States will report progress on measures included in the PIP Measurement Plan using PIP Progress Reports. States will also use progress reports to notify CB when the required amount of improvement is achieved. When this occurs, we will confirm that the required amount of improvement was attained and notify the state in writing of the outcome of our review and validation process. When the required amount of improvement is achieved, the state will no longer be required to report performance on that measure for the purposes of PIP measurement.

Attachment A. CFSR Round 4 Statewide Data Indicators Data Dictionary

Round 4 Changes

- As was done in Round 3, the trial home visit (THV) adjustment will continue to be applied when calculating performance on the indicator for permanency in 12 months for children entering foster care. In Round 4, we will apply adjustments only to those THVs that end 6 months after the 2 years needed to observe entries and exits. Round 3 used a 12-month, rather than 6-month, THV adjustment period. This change will reduce the number of AFCARS submissions needed to apply the THV adjustment from 6 to 5 and allow reporting of state performance on this indicator to occur 6 months earlier than in Round 3.
- For CFSR Round 4, we will use a new measure for reentry to foster care. The Round 3 indicator limited the population of children in the denominator to children who enter foster care during the 12-month reporting period and then discharged to permanency within 12 months of entry. The Round 3 indicator did not account for children with lengths of stay longer than 1 year and for whom the state has a responsibility to achieve permanency and prevent reentry to foster care. The new measure will not exclude children in care whose foster care episode(s) lasted less than 8 days. These changes will measure the desired outcome of discharging children to permanency and simultaneously minimizing reentry to foster care for all children in foster care regardless of length of stay in foster care.
- In Round 4, SWDI data quality checks will be performed only on records for the specified AFCARS 6-month submission period. In Round 3, checks were performed on all records included in the AFCARS submission, including records with activity dates prior to or after the 6-month submission period (e.g., entries or exits prior to or after the 6-month AFCARS period).
- Data quality checks pertaining to children in foster care at age 21 that were used in Round 3 will not be performed in Round 4 (i.e., age at entry greater than 21, age at discharge greater than 21, and in foster care more than 21 years).

Statewide Data Indicators: General Data Notes

- In cases where AFCARS records contain inconsistent information between two data submissions, the data elements from the latest submission are used.
- For the purposes of the CFSR indicators, children who turn 18 while in care are considered to have discharged from the foster care system on their 18th birthday and are not counted as achieving permanency.
- Children who enter care at age 18 or older, and who are age 18 or older on the first day of the 12-month period, are excluded from indicator calculations.

Title	Description	Denominator	Numerator	Risk Adjustment	Exclusions	Notes
Maltreatment in foster care	Of all children in foster care during a 12-month period, what was the rate of victimization, per 100,000 days of care?	Of children in care during a 12-month period, total number of days these children were in care as of the end of the 12-month period	Of children in care during a 12-month period, total number of substantiated or indicated reports of maltreatment (by any perpetrator) during a foster care episode within the 12-month period	• State • Age at entry (for children entering) or age on first day of the 12-month period (for children already in care)	• Complete foster care episodes lasting < 8 days are not counted in the number of days in foster care, and maltreatment that occurs within those episodes are not counted in the number of substantiated or indicated reports • Maltreatment reports that occur within the first 7 days of removal are not counted in the number of substantiated or indicated reports • Records with a victimization incident date outside the foster care episode, even if the report date falls within the episode (used when incident date exists)	• Cases are matched across AFCARS and NCANDS using AFCARS ID • Report date is used to determine if the victimization occurred in the applicable 12-month period regardless of disposition date. A record is included if the report date falls in the 12-month period, but the disposition date does not • Incident dates[3] are used (when reported) to help determine whether the victimization occurred outside the dates of the child's foster care episode • Date of prior discharge in the second 6-month file is used to calculate length of stay when: (a) the date of discharge is not reported in first 6-

[3] Victimization incident dates are not used to determine the occurrence or recurrence of child maltreatment for safety indicators.

Title	Description	Denominator	Numerator	Risk Adjustment	Exclusions	Notes
					• Records with a maltreatment report date outside the 12-month period • Subsequent maltreatment reports that occur within 1 day of the initial maltreatment report • Child foster care episodes where we cannot calculate the length of the episodes	month AFCARS file, (b) the date of prior discharge is reported in second 6-month file, *and* (c) the number of removals goes up by exactly one between the first and second 6-month file • Two years of NCANDS data are used to determine if a report of maltreatment occurred during the 12-month period as maltreatment reports are only included in the NCANDS Child File after the corresponding disposition is documented
Recurrence of maltreatment	Of all children who were victims of a substantiated or indicated maltreatment report during a 12-month period, what percent were	Number of children with at least one substantiated or indicated maltreatment report in a 12-month period	Number of children in the denominator who had another substantiated or indicated maltreatment report within 12 months of	• State • Age at initial victimization	• Subsequent victimizations with a report date that occurs within 14 days of the initial report • Subsequent reports in which the incident date matches the incident date in the initial report as they	• Report date is used to determine if the initial victimization occurred during the 12-month period regardless of disposition date • Use of incident dates is limited to determining whether maltreatment reports refer to same

Title	Description	Denominator	Numerator	Risk Adjustment	Exclusions	Notes
	victims of another substantiated or indicated maltreatment report within 12 months of the initial victimization?		their initial report		are treated as re-reports of same incident • Unborn children	incident, in which case, the duplicate record is excluded. The comparison of incident dates is a secondary check done only when incident dates are reported in the NCANDS file for each substantiated or indicated maltreatment report being compared
Permanency in 12 months for children entering care	Of children who enter care in a 12-month period, what percent discharged to permanency within 12 months of entering care?	Number of children who enter care in a 12-month period	Number of children in the denominator who discharged to permanency within 12 months of entering care	• State • Age at entry • State's foster care entry rate[4]	• Children in foster care < 8 days • Children who do not have a discharge reason • Children who re-enter foster care in the same reporting period due to AFCARS reporting structure not including	Trial home visit adjustment is applied[5]

[4] State's foster care entry rate is the number of children who enter care (the denominator for the indicator taking into account any exclusions) in a 12-month period divided by the State's census child population estimates as of July 1st of each year.

[5] Trial home visit adjustment: If a child discharges from foster care to reunification with parents or other caregivers after a placement setting of a trial home visit during any of the five data periods used for the indicator calculation, any time in that trial home visit setting that exceeds 30 days is discounted from the length of stay in foster care.

Title	Description	Denominator	Numerator	Risk Adjustment	Exclusions	Notes
					discharge reason for first episode	
Permanency in 12 months for children in care 12-23 months	Of all children in care on the first day of a 12-month period who had been in care continuously between 12 and 23 months, what percent discharged to permanency within 12 months of the first day?	Number of children in care on the first day of a 12-month period who had been in care continuously between 12 and 23 months	Number of children in the denominator who discharged to permanency within 12 months of the 1st day	• State • Age on first day	• Children who do not have a discharge reason • Children who re-enter foster care in the same reporting period due to AFCARS reporting structure not including discharge reason for first episode	
Permanency in 12 months for children in care 24 months or more	Of all children in care on the first day of a 12-month period who had been in care continuously for 24 months or more, what	Number of children in care on the first day of a 12-month period who had been in care continuously	Number of children in the denominator who discharged to permanency within 12 months of the 1st day	• State • Age on first day	• Children who do not have a discharge reason • Children who re-enter foster care in the same reporting period due to AFCARS reporting structure not including discharge reason for first episode	

Title	Description	Denominator	Numerator	Risk Adjustment	Exclusions	Notes
	percent discharged to permanency within 12 months of the first day?	for 24 months or more				
Reentry to foster care in 12 months	Of all children who exit foster care in a 12-month period to reunification, live with relative, or guardianship, what percent reentered care within 12 months of their discharge?	Number of children in a 12-month period who discharged to reunification, living with relative, or guardianship	Number of children in the denominator who reenter foster care within 12 months of their discharge	• State • Age at exit	• Children whose discharge from first episode did not have a discharge reason • Children whose date of latest removal for the reentry episode is reported as being before the date of discharge for the prior episode • Children who reenter foster care in the same reporting period due to AFCARS reporting structure not including discharge reason for first episode	If a child has multiple reentries within 12 months of their discharge, only the first reentry is selected
Placement stability	Of all children who enter care in a 12-month period, what was the	Of children who enter care in a 12-month period, number of	Of children who entered care during the 12-month period,	• State • Age at entry	• Children in foster care < 8 days • The initial removal from home (and into foster care) is not	When a child experiences multiple episodes in the same year and: (a) the date of discharge is missing in the first 6-month AFCARS

Title	Description	Denominator	Numerator	Risk Adjustment	Exclusions	Notes
	rate of placement moves per 1,000 days of foster care?	days[6] these children were in care as of the end of the 12-month period	number of placement moves during the 12-month period		counted as a placement move • Days in foster care for which the AFCARS file does not have placement information[7]	file and (b) the date of prior discharge in the second 6-month file is during the first 6-month file, then the date of prior discharge is used to calculate length of stay

[6] Days in care are counted in full day increment; thus children who enter foster care on one day and exit the next day are counted as being in foster care for 1 day.

[7] Since placement information is required to determine placement moves, days in foster care for which placement information cannot be determined are excluded from the count of total days in foster care.

Data Quality Checks: General Data Notes

- With few exceptions, individual records flagged by a data quality check are excluded from relevant indicator calculations; exceptions are included in the notes column.
- Data quality checks are only calculated on records with all the relevant values. In situations in which one of the relevant values is missing, the DQ check is not calculated. The exception to this is DQ checks designed to explicitly look for missing values.
- Data quality checks are performed only on records for the specified AFCARS submission period.

AFCARS Data Quality Checks

Title	Description	Denominator	Numerator	Limits	Notes
AFCARS IDs don't match from one period to next	Percent of records that do not match for a given record number in the next 6-month period	Number of children reported in first 6-month file	Number of children that do not match in the next 6-month file	> 40%	Individual records with this condition are not excluded from indicator calculations unless the overall percentage of non-matches for a state exceeds the DQ limit
Date of birth after date of entry	Percent of records where the date of birth is after the date of latest removal	Number of children reported in a 6-month file	Number of records where date of birth is after the date of latest removal	> 5%	
Date of birth after date of exit	Percent of records where the date of birth is after the date of discharge from most recent foster care episode	Number of children with a discharge reported in a 6-month file	Number of records where date of birth is after the date of discharge from most recent foster care episode	> 5%	
Dropped records	Record is missing a date of discharge, suggesting the child is still in care, but a record for	Number of children reported in first 6-month file	Number of children reported without discharge dates in first 6-month file	> 10%	Records that fail this DQ check are excluded from indicator calculations unless the failure only occurs in the last 6-month period of

AFCARS Data Quality Checks

Title	Description	Denominator	Numerator	Limits	Notes
	this same child in the next 6-month period does not exist		that do not appear in the subsequent 6-month file		the data periods[8] used in indicator calculations
Enters and exits foster care the same day	Percent of records where date of latest removal from home is the same day as the date of discharge from most recent foster care episode	Number of children with a discharge reported in a 6-month file	Number of records where date of latest removal from home is the same day as the date of discharge from most recent foster care episode	> 5%	
Exit date is prior to removal date	Percent of records where date of discharge from most recent foster care episode is before the date of latest removal from home	Number of children with a discharge reported in a 6-month file	Number of records where date of discharge from most recent foster care episode is before the date of latest removal from home	> 5%	
Missing date of birth	Percent of records with a missing date of birth	Number of children reported in a 6-month file	Number of children missing date of birth[9]	> 5%	
Missing date of latest removal	Percent of records with a missing date of latest removal	Number of children reported in a 6-month file	Number of children missing date of latest removal	> 5%	

[8] See state data profile for details on which data periods are used in performance calculations for each indicator.

[9] Date of birth is used to generate age in months for risk-adjustment when calculating a state's RSP.

AFCARS Data Quality Checks

Title	Description	Denominator	Numerator	Limits	Notes
Missing discharge reason (exit date exists)	Percent of records where date of discharge from most recent foster care episode exists but the reason for discharge is missing	Number of children with a discharge reported in a 6-month file	Number of records where date of discharge exists but discharge reason is missing	> 10%	
Missing number of placement settings	Percent of records with a missing number of placement settings	Number of children reported in a 6-month file	Number of children missing number of placement settings	> 5%	
Percent of children on first removal episode[10]	Percent of children where the total number of removals from home to date equals one	Number of children reported in a 6-month file	Number of records where total number of removals from home to date = 1	> 95%	Individual records with this condition are not excluded from indicator calculations unless the overall percentage of children on their first removal for a given state exceeds the DQ limit

[10] This data quality check helps to control for a state's ability to link historical experiences of children using the same AFCARS ID.

NCANDS Data Quality Checks

These checks are applied only to victims of maltreatment. A victim is a child for whom the state determined at least one maltreatment report for that child was substantiated or indicated. It does not include children receiving alternative response or unsubstantiated reports. It does include children who died, and the death was confirmed to be the result of child abuse and neglect.

Title	Description	Denominator	Numerator	Limits	Notes
Child IDs for victims match across years	Among victims, percent of unique NCANDS Child IDs that match across two consecutive NCANDS Child Files	Among victims, number of unique NCANDS Child IDs in either or both of two consecutive NCANDS Child Files	Of NCANDS Child IDs in the denominator, number of those IDs that appear in two consecutive NCANDS Child Files	< 1%	Individual records with this condition are not excluded from indicator calculations unless the overall percentage of matching child IDs is less than 1%
Child IDs for victims match across years, but dates of birth/age and sex do not match	Among victims, percent of unique Child IDs that match across years, but: (a) sex does not match, (b) date of birth does not match (only applicable to children under the age of one), or (c) age difference between years is outside of expected range	Among victims, number of unique NCANDS Child IDs that match across two consecutive NCANDS Child Files	Of NCANDS Child IDs in the denominator, number of those IDs with a different sex or an age difference between years that is outside of the expected range. For children under the age of 1, this also includes those IDs with differing dates of birth	> 5%	• An age difference outside of the expected range occurs when the child's age difference between the 2 years is less than 0 or greater than 3 • Children missing sex or age at report are not included in this DQ check Children under age 1 without DOB are excluded from this DQ check and related indicator calculations
Missing age for victims	Percent of victims with a missing age	Number of victims in NCANDS Child File	Of children in the denominator, number of children missing child age	> 5%	NCANDS validation process automates calculation of child victim age by subtracting DOB from report date, when DOB is included in the submission

Title	Description	Denominator	Numerator	Limits	Notes
Some victims should have AFCARS IDs in Child File	Percent of victims with an AFCARS ID in NCANDS Child File	Number of victims in NCANDS Child File	Of children in the denominator, number with an AFCARS ID in NCANDS Child File	< 1%	
Some victims with AFCARS IDs should match IDs in AFCARS Files	Some victims with AFCARS IDs should match IDs in AFCARS Files	Number of victims who have AFCARS IDs reported in NCANDS Child File	Of children in the denominator, number that match using AFCARS ID to a child record in AFCARS, during the same year	No	Individual records with this condition are not excluded from indicator calculations unless the state does not have at least one case with matching AFCARS IDs

Attachment B. Methodologies to Calculate and Categorize States' Risk-Standardized Performance Relative to National Performance

This section describes how we will calculate state risk-standardized performance (RSP) and categorize performance in comparison to national performance.

A. Calculating State Risk-Standardized Performance

State RSP on each statewide data indicator (SWDI) will be calculated using a multilevel (i.e. hierarchical) model designed for that indicator. A multilevel logistic regression model will be used for indicators in which the outcome for a child either occurred or did not occur (i.e., permanency in 12 months indicators, reentry to foster care, and recurrence of maltreatment). A multilevel Poisson regression model will be used for indicators in which the outcome is a count per unit of time (i.e., placement stability, maltreatment in foster care). We chose multilevel modeling because it is a widely accepted statistical method that enables accurate evaluation of relative performance among states with different case mixes. The multilevel models that we employ when assessing each state's performance take into account: (1) the variation across states in the age distribution of children served for all indicators, (2) the state's foster care entry rate for permanency in 12 months for children entering foster care, (3) the variation across states in the number of children they serve, and (4) the variation in child outcomes between states.

The result of this statistical modeling is a more accurate value and representation of each state's performance than can be obtained using the state's observed performance. We refer to this performance value as the state's RSP. It is similar to the statistical process used in healthcare to identify and adjust for variation in patient outcomes that stem from differences in patient characteristics (or risk factors) at the time of health care encounters. The RSP is the ratio of the number of predicted outcomes in the state over the number of expected outcomes (both obtained from the model), multiplied by the national observed performance. This ratio is similar to the "observed over expected" ratio used in other types of statistical analyses.

Risk Adjustment

This section describes risk adjustment and the key steps in the modeling approach, including the calculation of a state's predicted outcomes, expected outcomes, and RSP for the indicator.

Child Age

Risk adjustment is done on the child's age for each indicator (depending on the indicator, it is the child's age at entry, exit, first day of 12-month reporting period, or victimization). Adjusting on age allows us to statistically control for the fact that children of different ages have different likelihoods of experiencing the outcome, regardless of the quality of care a state provides.

State Foster Care Entry Rate

Risk adjustment is done on the state's foster care entry rate (per 1,000 children in the general child population) for one indicator: permanency in 12 months for children entering foster care. Adjusting on foster care entry rate allows us to control statistically for the fact that states with different entry rates have different likelihoods of risk for poor outcomes, regardless of the quality of care a state provides.

We calculate the foster care entry rate as the number of children entering foster care during the 12-month period divided by the number of children in the state's child population, multiplied by 1,000. We obtain the child population data from the population division of the U.S. Census Bureau. These census data reflect population estimates as of July 1 of each year, whereas the 12-month reporting periods we use to define children entering care are either October to September, or April to March. Therefore, we chose to use the census year closest to the 12-month reporting period the child entered foster care as the denominator. For example, if the indicator follows children who entered care between April 1, 2022, and March 31, 2023 (a "22B23A" file in AFCARS file conventions), we use child population estimates from the July 2022 census estimate. If the 12-month period spans October 1, 2022, through September 30, 2023, we use population estimates as of July 1, 2023.

Steps to Calculate Risk Standardized Performance

The process to calculate a state's RSP involves the following steps:

Step 1. Calculate each child's predicted probability for experiencing the outcome.

The regression model determines the predicted probability each child in the state will experience the outcome given his or her age and the state in which he or she resides. Probabilities range from 0.0 (0%, or outcome will never occur) to 1.0 (100%, or outcome will always occur). These probabilities reflect a child's given level of "risk" of experiencing the outcome. For example, the regression model might calculate that the expected probability of a 17-year-old in State X exiting to permanency within 12 months is .34, or 34%. For a 1-year-old in State X, the expected probability might be .54, or 54%.[11]

Step 2. Calculate the number of children in the state predicted to experience the outcome.

We sum the predicted probabilities for all children in the state to get the number of children we predict will experience the outcome. This predicted number is the number of outcomes (e.g., exits to permanency by 12 months, number of placement moves) we would predict the state to have based on the state's performance with its actual, observed case mix. Compared to the actual number of children who had the outcome, the predicted number is a better estimate of how the state is likely to perform assuming no change in case mix, policy, and practice.

Step 3. Calculate each child's expected probability for experiencing the outcome.

The same regression model then determines the expected probability that each child in the state will experience the outcome, given his or her age, if he or she were in the "average" state. It is similar to the

[11] A child's predicted probability is based on two values obtained from the model: (1) the risk associated with the child's age (i.e., the age-specific beta coefficient, or β) plus (2) the state's intercept, which reflects the underlying risk of experiencing the outcome in that state after accounting for the child's risk. The result of this sum is transformed to a probability to ease interpretation. The state's intercept is calculated based on the state's actual, observed performance relative to states with similar children—considering how many children it served, the age of these children, and how many of these children experienced the outcome. For permanency indicators, in which a higher number is more desirable, the intercept will be positive for a higher-than-average state, negative for a lower-than-average state, and close to zero for an average state. For the remaining indicators, the opposite is true. If there were no differences among states after adjusting for risk, the intercepts would be identical across all states.

predicted probability obtained in Step 1 but ignores the specific state the child is in. It does this by using data on all children in the nation and ignoring the state he or she is actually in. In other words, each age reflected in the nation's case mix has a given "risk," or probability, of achieving the outcome.[12] For example, the regression model might calculate that the expected probability of a 17-year-old discharging to permanency within 12 months of entering foster care, if he or she were in the "average" state, is .23, or 23%. For a 1-year-old, the expected probability might be .41, or 41%.

Step 4. Apply the expected probabilities.

The expected probabilities are applied to the children in each state: Infants in the state are assigned the probability of the outcome associated with being an infant in the "average" state; 17-year-olds in the state are assigned the probability for 17- year-olds in the "average" state; and so on.

Step 5. Calculate the number of children in the state expected to experience the outcome.

We sum the expected probabilities for all children in the state to get the number of outcomes we would expect if the state's children were served by the "average" state. This expected number is the number of outcomes expected if the "average" state had the state's same case mix.

Step 6. We take the ratio of the number of "predicted" outcomes over the number of "expected" outcomes.

For an indicator like permanency, where higher numbers are more desirable, a state with more predicted than expected exits can be said to have a higher-than-expected permanency rate; the state's ratio will be greater than 1 (e.g., 500 predicted / 400 expected = 1.25). A state with fewer predicted than expected permanency exits can be said to have a lower-than-expected permanency rate; the state's ratio will be less than 1 (e.g., 400 predicted / 500 expected = .80). A state with the same number of predicted as expected exits will have a ratio of 1 (e.g., 500 / 500 = 1.00), which suggests they perform no differently than the "average" state.

Step 7. Convert the ratio to a value with the same units as national performance.

To convert the ratio into a more meaningful value, we multiply it by the national performance. This puts the ratio into the same units (a percentage or rate per days in care) as the national performance. The result is the state's RSP. As a point of reference, a state with "average" performance (ratio = 1) will have an RSP equal to national performance (i.e., 1 x the national performance). All other RSPs will be better or worse than national performance.

Explanation of Risk-Standardized Performance

In sum, RSP is the ratio of the number of "predicted" outcomes over the number of "expected" outcomes, multiplied by national performance. RSPs can be compared to national performance and are relative to RSPs of other states. However, a state's RSP should not be compared to its observed performance. For example, if a state's RSP for recurrence of maltreatment is higher than its observed

[12] A child's expected probability is based on two values obtained from the model: (1) the risk associated with the child's age (i.e., the age-specific beta coefficient, or β) plus (2) the average intercept of all the states, which can be interpreted as the unique level of care the "average" state provides for its children.

performance, this does not mean the state's performance declined after risk adjustment. The converse is also true.[13]

Because national performance is essentially the "weighted mean" of the states' observed performance values, national performance will be pulled in the direction of states with larger case populations. The RSP ratio is multiplied by national performance to yield a value that is in the same units as national performance. This process of multiplying by national performance has no statistical properties in and of itself.

B. Categorizing State Performance Relative to National Performance

Because the states' RSPs and national performance are based on the same national case mix, a state's RSP can be compared directly to national performance to determine if the state performed better, worse, or no different than the "average" state.

To determine whether a state's RSP is statistically better, no different, or worse than national performance, we calculate 95% interval estimates around each state's RSP.[14] We will compare each state's interval estimate to the national performance[15] and assign each state to one of three groups:

- "No different than national performance" if the 95% interval estimate surrounding the state's RSP includes national performance.
- "Better than national performance" if the entire 95% interval estimate surrounding the state's RSP is better than national performance.
- "Worse than national performance" if the entire 95% interval estimate surrounding the state's RSP is worse than national performance.

[13] To determine the impact of risk adjustment, one strategy is to look at how a state's ranking changed before and after risk adjustment.

[14] The RSP is a complex function of parameter estimates and calculating exact interval estimates requires a computationally intensive bootstrapping process. Therefore, CB calculates approximate confidence intervals by using each child's beta coefficient, each state's intercept, the standard error of the intercept, and the traditional 1.96 multiplier. These confidence intervals produce results identical to those we obtained when using 95% confidence intervals around each state's intercept, which is a less communicable metric but a recognized approach to identifying groups that are statistically above or below a standard rate, like the national observed performance.

[15] Comparing the upper and lower interval estimates to the national performance is done using rounded versions of these data points. For indicators expressed as a percentage, the interval estimates and national performance are rounded to one decimal place. For indicators expressed as a rate (maltreatment in foster care and placement stability), the interval estimates are rounded to two decimals places. We chose to round these values because we do not believe it is appropriate to attempt to measure state performance to the degree of precision implied by three or more decimal places. The rate indicators are rounded to two decimal places due to the small rates associated with them.

Attachment C. Example Sustained Improvement Goals and High-Performance Values Across Baseline Performances for Sample Sizes of 33 and 65 Applicable Cases

Table C-1. Example Scaling Factors, Amounts of Improvement, Sustained Improvement Goals, High-Performance Values, and the Corresponding Number of Additional Strength Ratings Required for a Sample Size of 33 Applicable Cases for an Item

Baseline Performance	Strength Ratings	Scaling Factor	Amount of Improvement for Sustained Improvement Goal	Sustained Improvement Goal	Additional Strength Ratings Needed to Reach Sustained Improvement Goal	Amount of Improvement for High-Performance Value	High-Performance Value	Additional Strength Ratings Needed to Achieve High-Performance Value
0%	0	0.90	23%	**23%**	8	45%	**45%**	15
3%	1	0.87	22%	**25%**	8	43%	**47%**	15
6%	2	0.84	22%	**28%**	8	42%	**49%**	15
9%	3	0.81	21%	**30%**	7	40%	**50%**	14
12%	4	0.78	20%	**32%**	7	39%	**52%**	13
15%	5	0.75	19%	**34%**	7	37%	**53%**	13
18%	6	0.72	19%	**37%**	7	36%	**55%**	12
21%	7	0.69	18%	**39%**	6	34%	**56%**	12
24%	8	0.66	17%	**41%**	6	33%	**58%**	11
27%	9	0.63	16%	**43%**	6	31%	**59%**	11
30%	10	0.60	16%	**46%**	6	30%	**61%**	10
33%	11	0.57	15%	**48%**	5	28%	**62%**	10
36%	12	0.54	14%	**50%**	5	27%	**64%**	9
39%	13	0.51	14%	**53%**	5	25%	**65%**	9
42%	14	0.48	13%	**55%**	4	24%	**67%**	8
45%	15	0.45	12%	**57%**	4	22%	**68%**	8
48%	16	0.42	11%	**59%**	4	21%	**70%**	7
52%	17	0.38	10%	**62%**	4	19%	**71%**	7
55%	18	0.35	9%	**64%**	3	18%	**73%**	6
58%	19	0.32	8%	**66%**	3	16%	**74%**	6
61%	20	0.29	7%	**68%**	3	15%	**76%**	5

Baseline Performance	Strength Ratings	Scaling Factor	Amount of Improvement for Sustained Improvement Goal	Sustained Improvement Goal	Additional Strength Ratings Needed to Reach Sustained Improvement Goal	Amount of Improvement for High-Performance Value	High-Performance Value	Additional Strength Ratings Needed to Achieve High-Performance Value
64%	21	0.26	7%	**71%**	3	13%	**77%**	5
67%	22	0.23	6%	**73%**	2	12%	**79%**	4
70%	23	0.20	5%	**75%**	2	10%	**80%**	4
73%	24	0.17	5%	**78%**	2	9%	**82%**	3
76%	25	0.14	4%	**80%**	2	7%	**83%**	3
79%	26	0.11	3%	**82%**	1	6%	**85%**	2
82%	27	0.08	2%	**84%**	1	4%	**86%**	2
85%	28	0.05	2%	**87%**	1	3%	**88%**	1
88%	29	0.02	1%	**89%**	1	1%	**89%**	1
91%	30	N/A	N/A	N/A	N/A	N/A	N/A	N/A
94%	31	N/A	N/A	N/A	N/A	N/A	N/A	N/A
97%	32	N/A	N/A	N/A	N/A	N/A	N/A	N/A
100%	33	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Note: N/A indicates not applicable as ongoing measurement is not required.

Table C-2. Example Scaling Factors, Amounts of Improvement, Sustained Improvement Goals, High-Performance Values, and the Corresponding Number of Additional Strength Ratings Required for a Sample Size of 65 Applicable Cases for an Item

Baseline Performance	Strength Ratings	Scaling Factor	Amount of Improvement for Sustained Improvement Goal	Sustained Improvement Goal	Additional Strength Ratings Needed to Reach Sustained Improvement Goal	Amount of Improvement for High-Performance Value	High-Performance Value	Additional Strength Ratings Needed to Achieve High-Performance Value
0%	0	0.90	23%	**23%**	15	45%	**45%**	29
2%	1	0.88	22%	**24%**	15	44%	**46%**	29
3%	2	0.87	22%	**25%**	14	43%	**47%**	29
5%	3	0.85	21%	**26%**	14	43%	**48%**	28
6%	4	0.84	22%	**28%**	14	42%	**49%**	28
8%	5	0.82	21%	**29%**	14	41%	**49%**	27
9%	6	0.81	21%	**30%**	14	40%	**50%**	27
11%	7	0.79	20%	**31%**	13	40%	**51%**	26
12%	8	0.78	20%	**32%**	13	39%	**52%**	26
14%	9	0.76	19%	**33%**	13	38%	**52%**	25
15%	10	0.75	20%	**35%**	13	37%	**53%**	25
17%	11	0.73	19%	**36%**	13	37%	**54%**	24
18%	12	0.72	19%	**37%**	12	36%	**55%**	24
20%	13	0.70	18%	**38%**	12	35%	**55%**	23
22%	14	0.68	17%	**39%**	12	34%	**56%**	23
23%	15	0.67	17%	**40%**	11	33%	**57%**	22
25%	16	0.65	16%	**41%**	11	33%	**58%**	22
26%	17	0.64	17%	**43%**	11	32%	**59%**	22
28%	18	0.62	16%	**44%**	11	31%	**59%**	21
29%	19	0.61	16%	**45%**	10	30%	**60%**	20
31%	20	0.59	15%	**46%**	10	30%	**61%**	20
32%	21	0.58	15%	**47%**	10	29%	**62%**	19
34%	22	0.56	14%	**48%**	9	28%	**62%**	18
35%	23	0.55	15%	**50%**	10	27%	**63%**	18

Baseline Performance	Strength Ratings	Scaling Factor	Amount of Improvement for Sustained Improvement Goal	Sustained Improvement Goal	Additional Strength Ratings Needed to Reach Sustained Improvement Goal	Amount of Improvement for High-Performance Value	High-Performance Value	Additional Strength Ratings Needed to Achieve High-Performance Value
37%	24	0.53	14%	**51%**	9	27%	**64%**	18
38%	25	0.52	14%	**52%**	9	26%	**65%**	17
40%	26	0.50	13%	**53%**	9	25%	**65%**	16
42%	27	0.48	12%	**54%**	8	24%	**66%**	16
43%	28	0.47	12%	**55%**	8	23%	**67%**	16
45%	29	0.45	11%	**56%**	8	23%	**68%**	15
46%	30	0.44	12%	**58%**	8	22%	**69%**	15
48%	31	0.42	11%	**59%**	8	21%	**69%**	14
49%	32	0.41	11%	**60%**	7	20%	**70%**	14
51%	33	0.39	10%	**61%**	7	20%	**71%**	13
52%	34	0.38	10%	**62%**	6	19%	**72%**	13
54%	35	0.36	9%	**63%**	6	18%	**72%**	12
55%	36	0.35	10%	**65%**	6	17%	**73%**	12
57%	37	0.33	9%	**66%**	6	17%	**74%**	11
58%	38	0.32	9%	**67%**	6	16%	**75%**	11
60%	39	0.30	8%	**68%**	5	15%	**75%**	10
62%	40	0.28	7%	**69%**	5	14%	**76%**	10
63%	41	0.27	7%	**70%**	5	13%	**77%**	9
65%	42	0.25	6%	**71%**	4	13%	**78%**	9
66%	43	0.24	7%	**73%**	5	12%	**79%**	9
68%	44	0.22	6%	**74%**	4	11%	**79%**	8
69%	45	0.21	6%	**75%**	4	10%	**80%**	7
71%	46	0.19	5%	**76%**	4	10%	**81%**	7
72%	47	0.18	5%	**77%**	3	9%	**82%**	6
74%	48	0.16	4%	**78%**	3	8%	**82%**	5
75%	49	0.15	5%	**80%**	3	7%	**83%**	5

Baseline Performance	Strength Ratings	Scaling Factor	Amount of Improvement for Sustained Improvement Goal	Sustained Improvement Goal	Additional Strength Ratings Needed to Reach Sustained Improvement Goal	Amount of Improvement for High-Performance Value	High-Performance Value	Additional Strength Ratings Needed to Achieve High-Performance Value
77%	50	0.13	4%	**81%**	3	7%	**84%**	5
78%	51	0.12	4%	**82%**	2	6%	**85%**	4
80%	52	0.10	3%	**83%**	2	5%	**85%**	3
82%	53	0.08	2%	**84%**	2	4%	**86%**	3
83%	54	0.07	2%	**85%**	1	3%	**87%**	3
85%	55	0.05	1%	**86%**	1	3%	**88%**	2
86%	56	0.04	2%	**88%**	1	2%	**89%**	2
88%	57	0.02	1%	**89%**	1	1%	**89%**	1
89%	58	0.01	<1%	**90%**	1	<1%	**90%**	1
91%	59	N/A	N/A	N/A	N/A	N/A	N/A	N/A
92%	60	N/A	N/A	N/A	N/A	N/A	N/A	N/A
94%	61	N/A	N/A	N/A	N/A	N/A	N/A	N/A
95%	62	N/A	N/A	N/A	N/A	N/A	N/A	N/A
97%	63	N/A	N/A	N/A	N/A	N/A	N/A	N/A
98%	64	N/A	N/A	N/A	N/A	N/A	N/A	N/A
100%	65	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Note: N/A indicates not applicable as ongoing measurement is not required.