"Using Screening Data for Decision Making"

Slide 1: Welcome to the Webinar "Using Screening Data for Decision Making." This is one of 11 webinars developed by the staff at the National Center on Response to Intervention (NCRTI). This webinar focuses on the types of decisions that can be made using screening data at the district, school, grade, class, and individual level. Decision making is an essential part of any RTI model. Data-based decision making is part of the screening process, and data from the screening process are used for a variety of educational decisions. Data-based decision making is not just for student-level decisions. RTI requires entire systems (schools, districts, states) to use data-based procedures at all levels of implementation.

If you have not already, you may want to print out the accompanying handouts that go along with this presentation. My name is Dr. Rebecca Zumeta, and I will facilitate the next several slides. I am the coordinator of training, technical assistance, and product development for the NCRTI.

Slide 2: NCRTI has developed a series of webinars to cover information that is important for RTI implementation. On this slide, you can see that there are a variety of webinars available in this series. While you can watch the webinars in any sequence, we recommend that you first watch the "What Is RTI?" webinar in order to ensure you have an understanding of how RTI is discussed in this series. It is also recommended that you complete the "Understanding Types of Assessment Within an RTI Framework" and the "What Is Screening?" webinars prior viewing this webinar. Following this webinar, learn more about how to establish a screening process and select the best screening tool to fit your needs.

Slide 3: Upon completion of this webinar, participants should be able to understand and describe different types of decisions that can be made with screening data, learn how the placement of a cut score can influence who is identified as being at risk, and analyze different types of screening data.

Slide 4: One helpful way to check your understanding throughout the webinar is to complete the accompanying "Vocabulary" handout. The handout provides a table with a list of key terms on the left side, with columns to the right side for you to track your understanding before and after viewing the webinar.

Before viewing the webinar, you should make predictions of the meanings of each term. Then, throughout the webinar, you will complete the final meaning column, based on the definitions provided, along with space for a "Picture/Sketch/Example" column, where you can add anything helpful for remembering what you have learned. You can see how in this example, I was able to clarify the meaning of *primary prevention level* after watching the webinar.
If you have not already made predictions for the key terms of this webinar, please go ahead and pause the webinar do so now. Press "play" when you are ready to continue.

Slide 5: Screening is intended for more than just determining who may need an intervention. Decisions can be made at all levels—including district, school, and grade level. The next series of slides provides examples of decisions that can be made at each of these levels.

The data inform district-level decisions about support and allocation. District-level decisions related to screening include decisions about program improvement and curriculum decisions; innovation and sustainability decisions, general model effectiveness of implementation of a particular curriculum or RTI model; ensuring equitable services and supports across schools, access to supplemental supports, the ability to determine the effectiveness of instruction, and the determination of a specific learning disability. Allocation of resources and professional development may also be determinations that can be made with screening data.

Slide 6: School-level decisions related to screening may include identifying general school-level trends or issues and grade-level trends. Questions that may be asked are: Are there holes in grade-level performance? Do we lose subgroups of students at certain times?

You can also use screening data to evaluate the effectiveness of school-wide curriculum and instruction delivery. Is there consistent performance across grade levels? For example, data may indicate consistent growth from Grades 1–4 and then a significant drop in performance when students reach fifth grade. This may be an indication that fourth-grade instruction or curriculum is not adequately preparing students for fifth-grade expectations.

Finally you may use screening data to determine areas of need and guidance on how to set measurable school-wide goals

Slide 7: Screening data can also be used to make grade-level decisions. So you can evaluate data to determine if there are trends in performance of an entire grade level. You can determine the effectiveness of grade-level curriculum and instruction. You can use the data to determine areas of need and as a guide to set measurable grade-level goals by using initial screening data as a baseline measure or indicator of performance. You may also use screening data to identify students who may need additional instruction or assessment

Slide 8: A *cut score* is a score on a screening test that divides students who are considered potentially at risk from those who are considered not at risk. Cut scores might also be referred to as *cut-points* or *targets,* depending on the tool. It is important to remember that cut scores used to identify students as at risk may not be the same way you identify students who require supplemental support. This is referred to as the *target identification rate* and will be discussed in more detail later.

Slide 9: RTI success depends on accurate identification of the students who may be at risk. This is an essential component of preventing costly identification errors.

In a perfect world, screening would result in 100% accurate identification of students who were

truly at risk (or "true positives") and students that are truly not at risk (or "true negatives"). Cut scores are intended to differentiate those two groups of students that are at risk and students that are not at risk.

Cut scores are often set to overidentify students as at risk. This is because we would rather find more students that may be at risk and determine through follow-up assessment that they are not at risk than to miss the opportunity to intervene with a student that needs extra help. Your choice of cut score will affect classification accuracy of an assessment. As the second stage of the screening process, progress monitoring may be used to avoid inappropriately identifying students who do not need interventions, despite being identified as at risk through the screener.

Slide 10: Risk can be thought of in two ways. The first way is *categorical*, meaning the screener clearly indicates risk for a core outcome. For example, given the results of a particular test, you either have sickle cell anemia or you do not. The other way is a *continuous evaluation*. This means that there are not discrete categories for risk based on the screen. Instead, responses fall along a bell curve. For example, a student's performance falls along a bell curve where some kids will do well and others will not. However, a child's performance on the screener does not 100% accurately predict risk status. The cut score is the best guess based on data that is collected of who may or may not be at risk.

However, other factors may lead to the overidentification or underidentification of students. For example, a student may have been sick on test day, or a student is a "word caller" but lacks comprehension, or a student may have not been familiar with the examiner—which are all reasons that a student may look like they have lower or higher performance than they actually do.

The establishment of a logical cut score is essential for accurately identifying students for additional assessment and supports and for evaluating the effectiveness of the support system.

Slide 11: In this slide, we talk a bit about the clinical decision-making model. Again, in an ideal world, a screening tool would identify students at risk or not at risk with accuracy. Screening tools, however, are never 100% accurate because there are so many variables that can affect performance.

Screening tools vary in their sensitivity and specificity. *Sensitivity* is defined as the extent to which a screening measure accurately identifies students at risk for the outcome of interest. *Specificity* is the extent to which a screening measure accurately identifies students not at risk for the outcome of interest. Ideally, the screener would only indicate that a student is at risk if she is truly at risk. When students are correctly identified as at risk, it is called a "true positive." In addition, the hope is that the screener indicates that a student is not at risk if she is not at risk. When students are correctly identified as not at risk, it is called a "true negative."

Unfortunately screeners are not perfect. The screening tool might indicate that a student will be at risk when in fact he was not at risk. A case where the screener indicates that a student is at risk and the actual outcome shows that the student was not at risk, is called a "false positive." You

may also think of this as a false alarm. Alternatively, the screening tool might show that the student was not at risk when in fact he was. This is called a "false negative" because the screener indicated that he was not at risk when the outcome shows that he was at risk. In this case, students are incorrectly identified as not at risk with the screener.

Additional screenings and follow-up assessments are often used to verify the accuracy of initial screening results to help improve the extent to which students are correctly classified as at risk or not at risk.

Slide 12: Another consideration is the change in true positives, true negatives, false positives, and false negatives that occurs if you change the cut score. If you change the cut score, you will influence the number of students accurately identified or inaccurately identified.

In the example on the left, 80% of the poor readers are accurately predicted (meaning they were identified as at risk based on the screening instrument and were actually at risk based on the outcome) and 80% of the good readers are accurately predicted (meaning that the screening data showed them not to be at risk and outcome showed them to be not at risk). At the same time, the screener incorrectly identified about 20% of good readers and underidentified about 20% of poor readers.

In Example 2, the cut score was lowered. With the new cut score, about 65% of the poor readers were accurately predicted as at risk and about 5% of the good readers were incorrectly identified as at risk when they were not. So in other words, about 5% were false positive. By lowering our cut score, we misidentified approximately 35% of the poor readers and suggested that they were not at risk when they actually needed additional services.

Slide 13: So where should the cut score be set?

Typically, the best cut score is the one that produces the most accurate representation of who is at risk and who is not at risk, as shown with the center line. In this example, we are likely to identify students at risk as accurately as we identify students not at risk. Follow-up progress monitoring and screening will help increase the accuracy of the identification rate.

The cut score can be adjusted to meet the needs and priorities of schools or districts. They may ask themselves: Should the accuracy be increased by correctly identifying all students who are not at risk but accept the possibility of missing a larger group of students who are at risk? Or should the accuracy be increased by correctly identifying students at risk while increasing the number of students we may incorrectly identify (or what we call "false positives")?

Regardless of what cut score is used, it is critical that those interpreting the data understand the implications of the selected cut score. As noted previously, in most educational settings, we typically set the cut score so we overidentify students who are at risk because we would rather find students who may be at risk and discover that they are actually okay than to miss the opportunity to intervene because we set the cut score too low.

Slide 14: The risks associated with the over- and underidentification of students are similar to the risks found within a public health model.

In public health, overidentifcation can lead to higher expenses, resulting in additional testing and unnecessary worry. In education, overidentification can also lead to higher expenses related to testing as well as higher expenses related to early intervention services. In contrast, underidentification in public health can lead to missing a serious health problem; and in education, it can lead to a missed opportunity for prevention or early intervention.

To combat the costs associated with overidentification during universal screening, additional assessments—such as progress monitoring or diagnostic assessments—may be used to confirm or disconfirm the initial screening results. This is often referred to as the second stage of the screening process and can help reduce the amount of inappropriate intervention that may be provided to students that didn't need that support.

Slide 15: Establishing a cut score should be based on logical practices and be educationally relevant. Some screening tools include national norms to establish their cut scores. You could also use local norms as cut scores. Some data systems provide users with the tools to develop local cut scores based on outcomes on high-stakes tests in those users' particular state or district. Similar district assessment tools and cut scores allow for comparisons across schools. Often, data is needed in order to establish norms at the local level. You could also use cut scores based on the likelihood of demonstrated mastery on core testing or on standards based testing. This is also known as *benchmark testing*. A *benchmark* is a predetermined level of performance on a screening test that is considered representative of proficiency or mastery of a certain set of skills.

Cut scores are often considered arbitrary but are typically based on statistical analysis that allows us to determine the point at which the assessment will allow us to best classify students as at risk and not at risk based on our outcomes of interest.

Slide 16: There are some benefits for using district-established cut scores, as opposed to using school-established cut scores. District-established cut scores may lead to more effective and efficient allocation of resources and allow districts to target schools with greatest need, and identify effective approaches for scale up in other sites; it can allow district staff to make decisions about resources are based on data, and may allow those decisions to be more timely and relevant to the instruction that is occurring in schools. The district-based cut score can also increase buy-in and use of data by schools and teachers because district staff can clearly justify why the changes are being made. It also allows for a common message and focus around activities. It allows district staff to set more clear expectations across sites; and, by having a common message and focused activities, district staff are not only evaluating differences between schools but also changes across time. It may help staff to identify technical assistance and professional development activities that are based on demonstrated needs in sites, rather than perceptions or wants of staff. Finally, it has the potential to increase the equity in access to supplemental supports, so that the schools that are most in need of resources are able to access

them.

Slide 17: Let's look at three schools in the same district. These schools all established their own cut scores. As you can see from this table, School 2 appears to be doing the best and Schools 1 and 3 appear to have fewer students reaching the cut score. Numbers alone, however, can be deceiving. If districts allocate resources based on school-based cut scores, students in School 1 and School 3 would receive equal amounts of resources. In looking at this information, you must consider whether we are we comparing apples to apples.

Slide 18: Comparing the actual performance graphically, it is clear that School 1 significantly outperforms School 3. While School 1 only has 50% of students at or above the cut score, its cut score is higher than the cut scores for School 2 or School 3. Because each school established its own cut score, it is difficult for districts to determine the actual number of students at risk and make comparisons across schools.

Slide 19: Using a district-established cut score, however, it becomes clear that School 3 has a much lower percentage of students at or above the cut score (only 4%) than either School 1 (at 44%) or School 2 (at about 20%). School 3 would likely receive more intensive supports from the district because it demonstrates the greatest educational need. Even School 2, which had the highest percentage of students at or above the cut score in the school-based example, has only 20% at or above the district-established cut score. A standard cut score or target (as seen in this graph) across the district provides a more accurate picture of performance across the district and provides more accurate data for allocating district resources and targeting district supports.

Slide 20: NCRTI also recommends that teams establish routines and procedures for conducting data reviews, establish decision-making processes, and establish explicit decision rules for assessing student progress. This is an important task for district teams to accomplish so that there is some consistency across schools.

Slide 21: When conducting data reviews, it is important to develop routines and procedures for how your team will review data. Data teams should conduct data reviews at logical, predetermined intervals; should schedule data reviews prior to the beginning of instruction, to the extent possible; use established meeting structures or agendas; and to make sure you involve relevant team members.

Slide 22: It is important to have established routines and procedures for data-based decision making. In other words: What steps are you going to follow in order to use the screening data to make decisions?

Articulating routines and procedures in writing helps ensure and assess if established routines and procedures are being implemented with integrity. Ongoing evaluation of the selected routines and procedures is necessary to ensure they are culturally and linguistically responsive and lead to the desired outcome.

Slide 23: In thinking about establishing routines and procedures for data-based decision making, consider clarifying the following points in writing prior to implementation.

- **What are you looking for?** We all know that data "fishing" can be fun but may lead to problems. It can cause sites to delay the use of data, change the focus of the analysis, or miss important trends or issues. So it is important to identify what you are interested in knowing prior to your data analysis meeting. If you are unclear as to what you are looking for, conduct an analysis of the more critical outcomes first (i.e., graduation, reading performance) and then focus on outcomes in other areas. It is important to prioritize. Second, it is important to identify what you are looking for at all levels of analysis including the district, school, grade, class, and individual students and all levels of the prevention system.
- The second question you may ask is: **How will you look for it?** So develop a plan for how you will systematically analyze your data. This can increase the efficiency of your data analysis activities. It also helps manage the output many data systems offer. Only the most critical data are needed at first. It allows you to know where and when to delve deeper.
- The final question you might want to ask is: **How will you know if you found it?** Determine how much evidence is needed for the team to identify success or lack of success. Once identified, the team can continue moving through the problem-solving process in order to develop a plan of action.

Slide 24: Teams might also consider articulating—in writing—specific decision rules. They might want to define what happens when certain things occur, such as more than 80% of students being at or above cut score? What might they do if less than 80% have reached the cut score: Do we strengthen the core instruction and curriculum? Do we add an intervention for all students? If a lack of progress is evident: What should the team do? Do you make a plan? And it may be important to identify what will be done if student progress varies by target group (e.g., participation in Title I, receiving special education services, or have low socioeconomic status) .

Decision rules should be established at all levels, including class, grade, and school.

Slide 25: Data collection is not nearly as difficult for teams as data analysis. Screening data can assist teams in answering many of the questions they are asking. However, data analysis can be overwhelming.

This graphic is intended to demonstrate how overwhelming data analysis can be. Please keep in mind that the data presented in this presentation are fictitious but they are representative of data that may be collected and analyzed by a school or district. Many data systems provide users with more outputs than they know what to do with. Teams struggle with where to start, what data are most important, and what data provide the information they are looking for. Data systems can help manage information; however, if you are unfamiliar with the data system and what it has to offer, you may be paralyzed by its outputs. As a group, we will discuss some of the different

types of data analysis that can be done with screening data further in this section.

Slide 26: Data analysis should occur at the district level, school level, grade or class level, and individual student level.

Slide 27: The purpose of data analysis is to identify students who need additional assessment and instruction, evaluate effectiveness of core curriculum and instruction, allocate resources, and evaluate effectiveness of instruction programs for target groups of students.

Slide 28: This slide shows some of the more common terms that are sometimes confused while analyzing and interpreting data. You may have seen the terms used interchangeably, especially by publishers of screening tools. For purposes of this presentation, we refer to these terms as follows:

- A *cut score* is a score on a screening test that divides students who are considered potentially at risk from those who are considered not at risk.
- *Target*, or *benchmark*, is a predetermined level of performance on a screening test that is considered representative of proficiency or mastery of a certain set of skills.
- *Criterion scores* are scores on a screening test that separate students into performance levels (e.g., *established*, *emerging*, *deficient*)

The target or benchmark is, more often than not, the same score as the cut score—especially when the cut score is predictive of the state test. However, it does not have to be. For example, schools or districts may choose to set higher benchmarks or targets that are predictive of higher standards while setting cut scores to be more predictive of who is at risk and not at risk. This is why it is essential for teams to understand how the cut scores, targets, and benchmarks were established. The term *cut score* is sometimes used to refer to criterion scores that separate students into those performance levels—commonly depicted as the green, yellow, and red highlighted students. It is important to remember that the purposes of these two terms, *cut score* and *criterion score*, are different. We will talk more about this in a bit.

Slide 29: Before we begin analyzing some sample data, it is important to make sure everyone understands how screening data may be presented. Although we will be looking at fictitious data, the procedures we will be using can be applied to almost any type of data.

Screening data can be interpreted in several different ways. What you are interested in learning will determine what data you use (i.e., norm referenced, criterion referenced, or target scores) and how you choose to interpret your data. We will discuss each of these further.

Slide 30: If you are using norm-referenced data, Students are measured against others—not a defined criterion. It also permits a fixed proportion of students to pass and fail an assessment. This means that standards may vary from year to year, depending on the quality of the cohort, and it is an effective way of comparing students.

The norms used in norm-referenced data can occur at different levels. Students can be compared to other students in their class, school, or district, or against national or state norms. For example, the whisker plots shown within the next couple of slides provide examples of the comparison of students at the class, school, or district level. National norms and state norms are derived from formal norming studies.

Slide 31: Norm-referenced data compares an individual against his or her peers. A common way these data are represented is through a bell curve.

In the middle is the midpoint, median, or the 50% percentile. Half of the population is to one side of the line, and the other half is on the other side of the line. In a normal distribution, the majority of people fall near the middle, creating a perception of an average. The farther away from the middle, the fewer people you will find. In the education field, imagine the students who are very high achievers on the right end and very low achievers on the left end. In a typical classroom, it would be expected to have a similar distribution of skills among students, with very few high and very few low achievers, and the majority of students possessing similar skills.

Slide 32: Box plots are common ways to graphically present norm-referenced data. Think about the bell curve turned on its side. What you end up getting is a box plot. Box plots are somewhat similar in shape and representation. In a bell curve, the middle line is the median score, or 50th percentile; in the box plot, it is the middle line. The box represents the majority of the population, from the 25th to 75th percentiles, while the lines above and below represent the 75th to 90th percentiles and 10th to 25th percentiles, respectively. Students who fall above or below the box are considered to be outliers.

Box plots are helpful in comparing different groups or cohorts of students. They focus on where the average students fall, instead of the outliers.

Slide 33: This is an example of individual student data for a Grade 2 screening measure. This box plot compares an individual student to his peer group. The class or grade is represented by the box plot, and the student is represented by the dot. This level of analysis would be appropriate for identifying students in need of supplemental support.

As you can see in this example, the average score is about 45. Students in the 90th percentile scored 70; students in the 10th percentile scored about 15; and the score for the student that scored well below his or her peers was about 8.

In this example, you can see that the student is performing well below his peer group.

Slide 34: Norm-referenced data allows us to make comparisons of groups—in this case, a school compared to other schools. This type of analysis would be helpful in identifying which grades may be in need of additional support.

The green box plot provides a composite of all students in a state who have completed the same screening measure for that grade level in the fall. The grey box is School A's scores, and the black line is the target score. How does School A's first grade compare to the state's composite norm for first grade?

Although the median score for the state is below the target score, the majority of first-grade students in School A are below the target, suggesting that this school performs much lower than the norm. However, in Grade 2, the school is performing slightly better than the norm.

What about third-grade performance in comparison to their peers?

For Grade 3, the state norm is well below the target score. But even worse, all students in School A are performing below the target. It may be worth noting here that because most students—including those in the comparison group—are performing below the target, teams need to determine if larger systemic issues are affecting student performance. For example, they may need to consider whether they have ineffective policies, district curriculum, or teacher training.

Slide 35: Turn to the "Norm-Referenced Box and Whisker Plots" handout. As a member of a district-level data team, you are looking at the norm-referenced screening data for one school in your district, School A, compared to the composite for the state. What does the information in the graph tell you about School A? Pause the webinar to answer the questions on the handout.

We will now review possible answers to the questions; not all possible answers are represented.

- **What is the cut score?** The answer in this example is 60.
- **What is the 50th percentile performance for the composite?** The answer is approximately 75 for the composite and approximately 55 for School A
- **What is the spread or range of scores for the composite?** The range of scores is approximately 30–120; For School A, the range is 47–71.
- **What might the difference in spread between School A and the composite tell us?** The spread for School A is much narrower than for the composite. This may suggest the need for additional analysis to understand why the scores of School A are so limited. It might suggest that the teachers are teaching to the test rather than developing the overarching skills of the students.
- **What might you say about the performance of School A compared to the composite based on this graph?** You may consider that the median for School A falls below the cut score whereas the median for the composite falls above the cut score. This suggests that for the composite, more than 50% of the students are reaching the cut score; while at School A, less than 50% are achieving scores at or above the cut score.

As we just discussed, the distribution of scores for the composite is larger than for School A. This means that there is a larger range of proficiency between the strongest and weakest students

compared to school A.

So what additional questions might you ask based on these data? Some sample questions may be: Why is the distribution of scores so narrow for School A, or what supports are needed in order to help improve the performance of students at School A?

Slide 36: If you are using criterion-referenced data, you are able to evaluate students against defined or objective criteria. Criterion-referenced assessment is often, but not always, used to establish a person's competence or whether he or she can do particular skills or sets of skills. Criteria typically do not vary from year to year unless there is a change in what proficiency skills are required.

Norm-referenced data compares students to one another, but criterion-referenced data compares a student to a pre-set criterion.

Slide 37: There are multiple ways to determine the criterion. One example is percentile ranks. For example, below the 10th percentile may be an indication that the student may be in the *deficient* range; 10th percentile to 25th percentile may suggest the student is in the *emerging* range; and above the 25th percentile may suggest the student is in the *established* range.

Slide 38: This type of data representation is what is often associated with the triangle. In fact, if you turn this upside down, you will see where the triangle graphic originates.

Slide 39: The first table shows how students could be grouped if norm-referenced data is used. In this output, students are compared with their grade-level peers. Some students fall in the upper performance level, others in the lower performance level. This type of data output allows the grade-level teachers to see which students have the lowest and highest proficiency in the class and what the "average" performance group looks like. This can assist with grouping students for differentiated instruction.

The second table presents the same data using criterion scores to identify students. In this data output, you can see that all students are performing below the criteria. These data can assist in evaluating the efficacy of the core program.

Analyzing the same data in multiple ways allows for different questions to be answered. For that reason, different kinds of data sources can be useful for decision making in your school.

Slide 40: Data-based decisions can also be made using established benchmarks or target scores, not just the criterion-referenced scores. These scores are typically based on statistical analysis and can be correlated with high-stakes tests. Using target scores can assist with determining overall efficacy of core curriculum and instruction, while criterion-referenced scores can help with grouping of students for instructional programming.

Slide 41: The target score is typically a score where teams can feel confident that students have

mastered or are proficient on assessed skills. In other words, these are students the team is confident will not need additional assistance. In this example, several students who met the established criteria fall below the target score. Although they are demonstrating established skills, we are not as confident that they have mastered these skills and, thus, they will remain in the *established* range. Sites may consider progress-monitoring these students to ensure that they continue to make progress in the general education classroom.

Slide 42: Earlier we mentioned the district-level decisions that can be made using screening data. These include decisions about program improvement and curriculum, innovation, and sustainability, including the general effectiveness of implementation and general effectiveness of RTI model; it can also ensure equitable services and supports across schools and access to supplemental supports, access to effective instruction, and perhaps even initial indication for a student that may be at risk for a learning disability. Decisions about allocation of resources and professional development can also be assisted by evaluating data at the district level.

Slide 43: Districts could use data presented in this manner to compare schools at different grade levels and answer questions about effectiveness. For example: Is primary prevention working for most students in the schools? Districts can also use these data to make decisions about resource allocation. For example: Which schools appear to need additional support or further analysis?

Turn to the "District-Level Box and Whisker Graphs" handout. Pause the webinar to complete the questions.

We will now review some of the possible answers.

- **What does the information tell us about how primary prevention is working in schools in our district?** There is also a lot of variation in performance across the schools and the grade levels in this district. The median score for some schools in the district is above the target across all three grade levels and seems fairly consistent with the state comparisons, while other schools seem to be struggling or struggle in one grade. This variation may be especially worrisome to a district using a standardized district curriculum.
- **Which two schools in this district are struggling the most?** The answer to that question is School E and School A.
- **Which school in the district is doing the best?** School D.
- **What decisions might the district make about resource allocation? For example, which schools appear to need additional support or further analysis?** Looking at the data, the district may consider trying to determine what is happening at School E that causes it to perform lower across all three grades. It may be necessary to allocate additional resources to School E in order to bring the scores up. In schools where one grade level scores below the other grade levels, it may be necessary to do further analysis to see if additional resources need to be targeted at a specific grade level in order to improve scores.

Slide 44: Districts can also look at district-wide performance by grade level over the year. This chart shows changes in the percentage of students meeting the established criterion scores for each benchmark. This chart indicates that although only 55% of students met the criterion for *established* at the beginning of the year, 80% met it by the end of year. In general, it appears that instruction in Grade 2 is effective for most students. The district may be concerned, though, that the many students did not enter Grade 2 meeting the criteria for *established*.

Teams may ask themselves, what happened in first grade? Or why is second grade having to catch up?

Slide 45: Districts can also analyze performance by subgroup. Are students in Title I performing similarly to other students? Are students with disabilities performing at similar rates as students without disabilities? It appears, based on these data, that students in the "other" category, who are not identified with a specific status, are outperforming the target group in this district, yet both the Title I and special education populations fall below the target or benchmark scores score. It is also important to note that compared with other groups, students in special education are not making the same amount of growth across the year.

The district would need to obtain more information about special education students to determine why they are not closing the gap.

It's also important to look at the district-level data first to identify districtwide trends, before looking at school- and student-level data.

Not only are you able to see the differences in the performance of different groups in the spring, but you are also able to look at their growth rate or slope across the entire year.

Please note that in this example, "special education" refers to students identified as having a learning disability. Students in special education may still be receiving instruction within a general education setting.

Slide 46: Districts may also compare performance across ethnic groups. This graph shows that some ethnic groups are not performing at appropriate growth rates as others. Just as with the previous slide, you can look at the difference in growth rates or slope across different ethnic groups.

Slide 47: At the beginning of the school year, this district was able to see that there was a gap in performance between English language learners (ELLs) and non-ELLs. After analyzing for root causes and developing and implementing an action plan, the ELLs' performance improved over the course of the year. Although their performance does not match the performance as their peers in the school, the gap in their performance is closer to the target than it was at the beginning of the year.

As you can see, depending on how your screening data are coded, districts can use the data to analyze performance differences across a variety of factors, including socioeconomic status, grade level, curriculum, school schedule, or other areas of interest.

Slide 48: It is also important to consider school-level data analysis. You can use these data to identify school trends or issues. Are there holes in grade-level performance? Do we lose subgroups of students at certain times? We can use the data to evaluate the effectiveness of school-wide curriculum and determine areas of need and guidance and use the data to create measurable school-wide goals.

Slide 49: Schools can analyze performance across grade levels. This graph indicates that Grade 2 and Grade 3 may need additional support.

A school with this data would identify general trends and issues, analyze for causes, develop a plan, implement and monitor the plan, and then evaluate and adjust the plan based on the ongoing data that they collect. These data could also be used to help to guide basic resource allocation across grade levels.

Slide 50: Schools can look at growth by benchmark period and by grade level by looking at the performance of an average student across different grade levels. The average student appears to be making growth at all grade levels, with more growth in the earlier grades in this example.

Slide 51: Schools, like districts, can also look at growth by target group. Turn to the "Analyzing Growth of Ethnic Groups at the School Level" handout. Pause the webinar and use the data provided to answer the questions.

We will now review some of the possible answers.

- **Which ethnic groups are performing above the target score in this school?** The answers are Asian, Caucasian, and Unidentified categories.
- **Which ethnic groups are performing below the target score?** The answers in this example are African American and Hispanic.
- **Consider the growth of students by ethnic group from fall to spring. What does this tell you about the achievement gap between ethnic groups?** While there is a difference in the fall between the students identifying as Asian, Caucasian, and Unidentified and those identifying as African American and Hispanic, the widening spread by the spring suggests that the achievement gap is growing across the school year. This is a cause for additional concern because the rate of the improvement for some ethnic groups is outpacing that of other ethnic groups. Asian, Caucasian, and Unidentified students are making more growth across the year than African American and Hispanic students in this example; that means that the gap is widening over time.

If these data represented your school, what next steps might your team consider?

Slide 52: At the grade and class level, you can look at grade-level or class-level trends, effectiveness of grade-level curriculum and instruction, areas of need, and guidance on how to set measurable grade-level goals, and to identify students who may need additional instruction or assessment.

When analyzing data at the grade level, it is important to remember that the focus is on grade-level data and not individual students. Look at the big picture first. It should be an efficient and systematic process. Remember your model may look different to make it more efficient in your context. Data analysis procedures should be efficient, systematic processes that have been pre-established for your team. Analysis must guide and inform core instructional decisions.

Grade-level data analysis should not be about compliance, but rather about improving efficiency and instruction.

Slide 53: Grade- and class-level teams can also look at screening data to see the effectiveness of the core curriculum and how changes to the core instruction influences students' responses across the school year. This is an example of grade-level data. A similar process could be used for class level data.

Turn to the "Grade Level—Analyzing Effects of Changes to Instruction" handout. Pause the Webinar to answer the questions based on the data provided.

We will now review some of the possible answers.

- **Overall, how are the students in second grade doing in the spring? Has this changed across the school year?** The data indicate that the change led to improvements in the number of students meeting the *established* criteria. Between the winter and the spring, the percentage of students meeting the *established* criteria increased from 53% to 76%. The change in instruction led to an improvement from the winter to the spring, but the growth across the entire year was minimal, from only 73% to 76%.
- **What percentage of students requires tertiary prevention during the spring? What questions might you ask about this?** About 10% of the students likely need tertiary prevention during the spring. This is twice the typical recommendation of 5 percent. Across the entire year, there has been a high percentage of students being identified as *deficient*, or requiring that tertiary instruction. This is something that might cause concern and be in need of further analysis by the school RTI team.
- **How might this be different if you were looking at district-level data or school-level data?** The schools or districts might consider how the distribution shows a need for additional resources or supports in order to address the low level of students responding to the instruction in the winter. They might consider making curriculum changes or implementing additional professional development to support teachers in their implementation of the core curriculum.

Slide 54: Growth, just like at the school and district, can be analyzed at the grade and class level. Given the large number of data (i.e., number of students in the analysis) at the district and school levels, some gaps at the grade or class level may be missed. Class and grade levels should follow the same data-analysis routine to identify potential gaps in performance. Just as with the district-level analysis, you can look at the data by target groups, ethnic groups, or ELL status. In this class, the general education students are making progress above the target groups' benchmark performance, but the special education students remain below the target group and show little progress. In this graph, you can also see the difference in growth by the different subgroups in the class. Students receiving special education services have made almost no growth across the year, indicating that additional supports may be necessary for those students.

Slide 55: This chart shows average student performance for four fictitious classrooms. Teams looking at classroom data can compare student performance across classrooms at different points in the year.

Teams can also look at the data to make decisions about resource allocation, instruction, and evaluation of the core curriculum—not to evaluate teachers. There should not be fear about the data among teachers and school teams. The discussion should be open.

The graph shows that in the fall, the classes were performing at different levels. The school may need to allocate resources appropriately to ensure that Classes 1, 3, and 4 can close the gap.

It looks like Class 1 made a lot of gains. This may be a result of additional resources, fidelity of implementation, additional teacher training, or other variables. Further analysis is needed to determine and address performance differences.

Spring data indicate that, on average, students in Classes 1, 3, and 4 are not performing similarly to Class 2. Adjustments may need to be made in the following grade to account for differences. It may also be worthwhile for the teaching team to evaluate the consistency in which they are delivering instruction across the classes in their grade level.

Slide 56: Ranking graphs can also be used to conduct student-level analysis at the grade or class level. This type of analysis can allow you to see how individual students rank against other students in the class or grade. Teachers can use this information for differentiating instruction or to create small instructional groups. For example, you may group the students shown in the black boxes together in a small group.

Slide 57: When class- and grade-level analysis is complete, data teams may consider looking at the performance of individual students to identify which students need additional support in comparison with their peers, the target score, or other performance criteria.

Slide 58: With screening data, individual student's performance can be compared with that of their peers. If you look at the green boxes, you can see that the group appears to be making

progress over time. While this student shown by the blue dot is making some progress across the year, his progress appears to be insufficient to close the gap with his peers. In fact, if you look closely, it appears the gap may actually be increasing, as shown by the increasing difference between the 50th percentile of the peer group and the student's benchmark score. This may be a student who requires additional support.

Slide 59: Analyzing student data are important—not just for students who are struggling. It is important to also ensure that high performers continue to make the appropriate levels of progress throughout the school year. This may also be a student who is in need of additional support.

Slide 60: How is screening data used to determine who will need supplemental support? After determining the effectiveness of primary prevention and establishing instructional plans for improving core instruction, teams can begin to analyze data to determine which students need additional support.

Identifying students in need of additional support is not always as simple as providing interventions for all students well above or well below the cut score or target score. It may vary based on the needs and resources of the school and the target or criterion scores chosen.

There may be more students below the target score than can be served by the existing resources. Schools and districts may need to identify the lowest percentage of students whose needs can be met using existing resources (i.e., 20%), and they may need to think about reallocating existing resources or securing additional funding to increase the number of students they can serve in interventions.

Remember, if there are many students below the target (i.e., more than 20%), there has to be some changes to the primary level or core instruction so the focus should be on improving core instruction/curriculum.

Slide 61: Practitioners can use criterion scores from criterion-referenced assessments to determine what level of instructional supports students may need. In this example, students in the yellow may need secondary support while students in the red may need tertiary support. Although this method does provide established scores for grouping, it can lead to challenges for schools where a large number of students fall into these categories.

Slide 62: This graph provides an example of a situation where most students fall below the cut score or target score. If cut scores or target scores were used to determine supplemental intervention in this district, then School 3 would have significantly more students to support. This would be an issue of inequity in terms of resources. The focus for all of these schools, given the high numbers of students in need across buildings, should be on improving core instruction and curriculum in addition to providing interventions to students with significant needs.

Slide 63: Another way to identify students in need of support is to use the target identification rates. This is different from the cut score or target score. The *target identification rate* is the

proportion of students to be identified as at risk. This may depend on the program objectives and resources. There might be unique target identification rates specified for different skill areas. For example, a school might decide to select a larger target identification rate for word reading than for comprehension, due to resource availability. Schools and districts will also need to think about reallocating resources or securing additional funds to support *all* students in need.

It is important to emphasize that using a target identification rate to identify students at risk does not mean that schools are not responsible for teaching all students. Ideally, schools should provide extra supports to all students who demonstrate a need. Setting a target identification rate can be a first step while schools are (a) strengthening their core curriculum to reduce the overall numbers of students identified as at risk, and (b) reallocating resources or securing additional funding to support students who need additional interventions and supports.

Slide 64: This is an example of how the target identification rate may differ from school to school. School 1 has resources available to serve 20 percent of the students in secondary or tertiary intervention levels; therefore, it sets its target identification rate as identifying the lowest 20 percent of students in requiring interventions. In contrast, School 2 only has enough resources available to serve 15 percent of students in secondary and tertiary, so its target identification rate is set to identify the lowest 15 percent.

Regardless of a school's target identification rate, if more than 20 percent of the student population is identified as at risk, the focus should be on improving core curriculum and instruction.

The district should help identify resources available and help guide schools in allocating resources to support all students who need additional supports.

Slide 65: When making data-based decisions, it is important to remember these things. First, good data *in* means good data *out*. Know where your data came from and the validity of that data. Focus on the big picture—or all students. Are most students making progress? If not, what can we do differently to improve outcomes for all students. All instructional and curriculum decisions should be based on data. Keep data collection and analysis procedures simple and efficient, and teachers and leaders need to be able to trust the data.

When working with data, consider these questions: Is there a simpler way to do this? Do we need these forms or processes? Are we getting to the point?

Slide 66: Thank you for taking the time to listen to "Using Screening Data for Decision Making." To find more resources on this topic or to view other webinars in the implementer series, visit www.rti4success.org.You can also view additional information from the RTI Action Network and the IDEA Partnership.