REL WEST | WestEd

District Guide for Creating Indicators for Early Warning Systems

District Guide for Creating Indicators for Early Warning Systems

Prepared by American Institutes for Research for REL West at WestEd

Yibing Li
Jenny Scala
Dean Gerdeman
David Blumenthal

About REL West

The Regional Educational Laboratory West (REL West) at WestEd, serving Arizona, California, Nevada, and Utah, is part of a national network of 10 RELs whose mission is to provide research, analytic support, and resources that increase the use of high-quality data and evidence in education decision-making. Most REL West work is carried out in partnership with educators — from state and local decision-makers to district and school support providers and practitioners — through eight regional research alliances. For more information, visit http://relwest.WestEd.org.

This resource was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0002 by Regional Educational Laboratory (REL) West at WestEd. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Suggested citation

This publication and ancillary materials are in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

Li, Y., Scala, J., Gerdeman, D., & Blumenthal, D. (2016). *District Guide for Creating Indicators for Early Warning Systems*. San Francisco: REL West @ WestEd.

These materials are available on the REL West website at https://relwest.wested.org/resources/234.

Table of Contents

Introduction 1
- Background 1
- Audience 2
- Organization of the Guide 4

Section 1. Review Commonly Used Early Warning Indicators 5
- What to Expect in Section 1 5
- Step 1.1 Review The Research About Early Warning Indicators 6
  - Attendance 6
  - Behavior 6
  - Course Performance 6
  - Composite Indicators 6
  - Other Indicators 7
- Step 1.2 Consider What Types of Indicators and Cut Points Could Apply in Your District 7
- Intended Products 8

Section 2. Select Student-Level Data and Construct Longitudinal Cohorts 9
- What to Expect in Section 2 9
- Step 2.1 Identify Student Cohorts to Examine 10
  - Select outcome of interest 10
  - Determine the grade levels of focus 10
  - Identify the student cohorts 10
- Step 2.2 Create a Crosswalk of Potential Indicators and Local Data Sources 11
- Step 2.3 Prepare a Dataset for Analysis 12
- Intended Products 14

Section 3. Identify Trends and Patterns Found in Graduation Rates and Indicators 15
- What to Expect in Section 3 15
- Step 3.1 Check the Distribution of the Outcome Variable 15
- Step 3.2 Review the Distribution of the Predictor Variables Using Frequency Tables and Visual Tools 16
- Intended Products 19

Section 4. Analyze Relationships Between Potential Predictors and Graduation Outcomes 21
What to Expect in Section 4 21
Step 4.1 Summarize Potential Predictors for Students Who Achieved the Graduation Outcome and Those Who Did Not 22
Step 4.2 Visually Depict Relationships Between Potential Predictors and the Outcome 23
Step 4.3 Identify Predictors with Meaningful Differences Across the Two Outcome Categories 24
Intended Products 25

Section 5. Choose the Best Cut Point for Early Warning Indicators 27
What to Expect in Section 5 27
Step 5.2 Check the Performance of the Indicators Against the Actual Outcome 28
Step 5.3 Compare Classification Performance Based on Common Measures 29
Intended Products 30

Section 6. Test the Relative Performance of Early Warning Indicators 33
What to Expect in Section 6 33
Step 6.1 Choose a Criterion That Is Aligned With the Local Priorities 33
Step 6.2 Compare the Indicators and Actual Student Outcome 35
Step 6.3 Decide the Number of Indicators or the Combination of Indicators 35
Step 6.4 Decide the Level of Risk on Which the System Should Focus 35
Intended Products 36

Section 7. Examine the Reliability and Accuracy of Different Indicators or Combinations of Indicators 37
What to Expect in Section 7 37
Step 7.2 Apply the Indicators to a Different Cohort of Students 38
Step 7.3 Refine and Reflect on Indicators 38
Intended Products 38

Conclusion 39

References 40

Appendices 43

Appendix A. Early Warning Indicator Documentation Template (from Section 1) 43

Appendix B. Map of Data Availability (from Section 2) 45

Appendix C. Data Quality Checklist (from Section 2) 46

Appendix D. Documenting the Distribution of the Outcome Variable (from Section 3) 46

Appendix E. Graduation Outcomes Tracker (from Section 4) 48

Appendix F. Template for Calculating Performance Measures of Indicators and Identifying Optimal Cut Points Through Receiver Operating Characteristic Analysis (from Section 5) 49

Identifying Optimal Cut Points Using ROC Curve Analysis 51

Appendix G. Indicator Criterion Decision Matrix (from Section 6) 52

Introduction

Background

Across the country, many K–12 school districts have implemented early warning systems intended to improve education outcomes for students. These systems use readily available data to identify students who are at risk of failing to achieve a desired outcome, such as graduating from high school on time. By identifying at-risk students, an early warning system allows schools to proactively support them so they achieve the desired outcome, rather than rely on reactive, remedial interventions.

The identification process at the core of an early warning system relies on constructing indicators that accurately predict students' future education outcomes. Early warning indicators can be thought of as predictive measures of students' likelihood of achieving outcomes, based on such information as students' attendance history, disciplinary infractions, and academic performance. Research has shown that the predictive strength of indicators and the thresholds or cut points for accurately identifying at-risk students may vary across contexts, underscoring the salience of the local experience when building early warning systems (Allensworth & Easton, 2005; Hartman, Wilkins, Gregory, Gould, & D'Souza, 2011; Norbury, Wong, Wan, Reese, Dhillon, & Gerdeman, 2012; Stuit, O'Cummings, Norbury, Heppen, Dhillon, Lindsay, & Zhu, 2016).

Currently, limited guidance exists on how individual districts can make data-based choices about what local indicators are appropriate to use for their own early warning systems. This guide provides practical advice to help districts explore their own student data, to identify student attributes and behaviors that are related to education outcomes, and to make informed decisions in developing or refining accurate early warning indicators.

Implementing an early warning system has multiple components, which include trained staff who can interpret and act on the data, along with a system for assigning appropriate supports and monitoring their effectiveness. This guide, however, focuses on the foundation for an early warning system for dropout prevention: determining which locally informed indicators best predict students at risk of not achieving desired graduation outcomes, then constructing those indicators to match the local need.

Research has shown that the predictive strength of indicators may vary across contexts, underscoring the salience of the local experience.

Audience

This guide is intended primarily for school districts that have experience with early warning systems and are interested in refining or expanding their own early warning indicators.

The content of the guide will be most useful for audiences that are knowledgeable about student data systems and are capable of working with spreadsheets or statistical analysis software. The guide provides general advice on data and analytic techniques that districts can consider in developing and refining locally tailored early warning indicators. The steps contained in each section should be considered as guideposts to be adapted to match local capacity and data constraints.

For readers who wish to learn more about the process of implementing an early warning system, including establishing teams and mapping appropriate interventions to individual student needs, companion publications from the network of Regional Educational Laboratories and other federally funded centers delve deeper into these topics and are available at the links below.[1]

1 Frazelle, S., & Nagel, A. (2015). *A practitioner's guide to implementing early warning systems* (REL 2015–056). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. Retrieved from http://ies.ed.gov/ncee/edlabs.

Therriault, S. B., O'Cummings, M., Heppen, J., Yerhot, L., Scala, J., & Perry, M. (2013). *Middle grades early warning intervention monitoring system implementation guide.* National High School Center. Retrieved July 20, 2016, from http://www.earlywarningsystems.org/wp-content/uploads/documents/EWSHSImplementationguide2013.pdf.

Key Terms Used in This Guide

The following terms are often used in discussing early warning systems and indicators. All descriptions provided assume application within an early warning system.

- **Cohort:** In a school context, this term refers to students who start an education program at the same time. A grade level or graduating class of students are the most common examples. A cohort of students may begin an education program (e.g., grade 9) at the same time, but individual members might progress through the program on different trajectories and might not complete the program with the rest of their cohort.
- **Cut point:** In the context of an early warning indicator, the cut point for an indicator is the specific value above or below which students are flagged as being at risk. For example, if a student is flagged as being at risk once he or she has missed 10 percent or more of instructional time, then 10 percent is the cut point for an indicator related to instructional time. Other common terms used when referring to cut points are thresholds, benchmarks, and cut scores.
- **Dropout:** A term for a student who did not achieve a high school graduation outcome. Districts must choose a precise, consistent definition of what it means for a student to drop out. (The term dropout is also used to identify the education outcome for these students.)
- **Early warning system:** A process that relies on readily available data to identify, support, and monitor students at risk of failing to achieve desired outcomes, such as on-time graduation. Early warning systems include, at minimum, (1) an outcome of interest determined by the district, and (2) several indicators shown to predict students' likelihood of achieving the outcome.
- **Flag or flagged:** Students who fall within a designated risk range, either below or above a cut point depending on the indicator, are referred to as receiving a flag, or being flagged, indicating that the student is at risk.
- **Indicator:** An early warning indicator is a measure of student characteristics that fall above or below an assigned cut point associated with high risk of not achieving the desired outcome. For example, a value of 1 on a course failure indicator may be defined as failing at least one course and a value of 0 on the same indicator may be defined as passing all courses. Receiving a score of 1 is an indicator that the student may be at risk. Indicators are usually derived from predictors (defined below).
- **Outcome:** In the context of an early warning system, an outcome refers to the result the system is intended to track. This is often reported at the student level as a binary variable: did a student graduate on time (yes or no); did a student graduate within five years (yes or no); did a student drop out (yes or no); and so forth. At the school or district level, the outcome may also be aggregated as a percentage of the whole (e.g., "85 percent of the 2012 cohort graduated on time").
- **Predictor:** A quantifiable measure of student performance or engagement, such as attendance rate, number of suspensions, or grade point average (GPA). A predictor could be continuous, such as attendance rates, or categorical, such as the number of suspensions, or binary, such as "failing a course." In this example, an attendance rate could fall anywhere along a 0 to 100 scale, whereas number of suspensions falls into discrete categories, such as 1, 2, or 5. Predictors can be turned into indicators by assigning cut scores around values associated with increased risk of not achieving the desired outcome.

continued ≫

- **Student-level data or records:** Student-specific data or records include information on an individual student, such as student identification number (student ID), year the student entered grade 6 or grade 9, number of absences, GPA, suspensions, and so on. Each row in the data represents the unique student and contains all the values associated with that student. If data exist spanning multiple years, each row could represent a unique student-school year combination (for example, a row containing all data for the student from the 13-14 school year, and another row containing all data for the same student from the 14-15 school year). Sometimes referred to as *administrative data* when collected centrally by a school or district.
- **Subgroup:** A term to describe a smaller group of students who share some common characteristics from within the larger student sample or population. This subgroup may share a common characteristic, such as being members of a particular ethnic group, or a combination of characteristics, such as being female English language learners.
- **Variable:** A term for data that varies between students and that can be used to create subsets of students. A variable may be a student ID, school year, or quantifiable measure of student performance or engagement, such as attendance rate, number of suspensions, or GPA. In a file containing student data, each variable is assigned its own column.

Organization of the Guide

This guide is organized into seven sections:

- Section 1 reviews commonly used early warning indicators and summarizes the research behind these indicators.
- Section 2 explains the process to select a student sample, explore the local data sources, and prepare a dataset for analysis.
- Section 3 discusses simple descriptive ways to understand the distributions of potential predictor and outcome variables with the local context.
- Section 4 provides examples of descriptive ways to understand the relationships between potential indicators and graduation outcomes.
- Section 5 discusses the steps needed to establish appropriate cut points for early warning indicators.
- Section 6 explains the rationale and steps for comparing early warning indicators.
- Section 7 discusses how to examine the reliability and validity of indicators or a combination of indicators.

The sections guide users through the technical steps involved in developing or refining early warning indicators that reflect their own local context. Tables, graphs, and templates are included as practical examples. The templates are available electronically for customization in a companion Word document: *District Guide for Creating Early Warning Indicators — Editable Planning Forms*.

Section 1. Review Commonly Used Early Warning Indicators

What to Expect in Section 1

By the end of this section, you will be able to…

- Review commonly used early warning indicators that have been shown to reliably predict graduation outcomes (Step 1.1)
- Consider the relevance of commonly used indicators and cut points to your district (Step 1.2)

Use Appendix A to summarize indicators of interest or currently in use locally. The appendix contains a template that groups the indicators into attendance, behavior, course performance, and composite categories.

Step 1.1 Review The Research About Early Warning Indicators

Early warning indicators are typically based on student attributes and behaviors that are malleable (i.e., can be influenced by school intervention or policy) and are related to an outcome of interest, such as on-time graduation. The existing literature on dropout prevention provides a strong starting point to guide the decision-making process for developing effective indicators.

The early warning indicator research identifies as many as 110 distinct indicators that have been used to predict which students are at risk of failing to complete high school (Bowers, Sprott, & Taff, 2013). The majority of these indicators have been constructed based on readily available data on "the ABC's" of student attendance, behavior, and course performance (Allensworth & Easton, 2007; Balfanz, Herzog, & Mac Iver, 2007; Frazelle, Negel, & Northwest, 2015; Roderick, 1993). The following paragraphs break down how these three broad families of predictors — attendance, behavior, and course performance — have been constructed into specific indicators in the early warning system literature. You may find it beneficial to review some of this research directly.

Attendance

The number of days students are absent is one of the most practical predictors for identifying students who are in need of interventions (Allensworth & Easton, 2007). Attendance variables that have been used as early warning indicators include overall semester or yearly attendance, daily attendance over shorter intervals such as the first 20 days in a school, course-by-course absences, and daily absences (Allensworth & Easton, 2007; Heppen & Therriault, 2008; Uekawa, Merola, Fernandez, & Porowski, 2010).

Behavior

Common behavioral indicators of dropout risk include office discipline referrals, suspensions, detentions, or classroom behavior grades or marks (Balfanz, Herzog, & Mac Iver, 2007; Uekawa et al., 2010). Compared with attendance and course performance, behaviors are used less frequently to develop early warning indicators due to inconsistencies commonly found with how disciplinary data are reported and the lack of behavior grade data.

Course Performance

Poor course performance is one of the most consistent predictors of dropout, whether measured through grades, test scores, credit accumulation, or course failure (Alexander, Entwisle, & Kabbani, 2001; Battin-Pearson, et al., 2000). Districts usually maintain three types of course performance data: (a) course failures, particularly in core academic courses; (b) GPA; and (c) number of credits accrued in each term or by end of school year (Allensworth & Easton, 2005, 2007; Balfanz et al., 2007; Neild & Balfanz, 2006).

Composite Indicators

Composite early warning indicators combine two or more predictors. Two prominent examples that combine multiple course performance measures include the Total Quality Credit indicator and the On-Track Indicator. The Total Quality Credit is a combination of course grades in core subject areas (Carl, Richardson, Cheng, Kim, & Meyer, 2013). The On-Track Indicator, developed by the University of Chicago Consortium on Chicago School Research, is based on accumulation of course credits and failures in core subject-area courses in grade 9 (Allensworth & Easton, 2005). Other researchers have found these two composite indicators to be simple and highly predictive indicators that can be applied in many contexts (Knowles, 2015).

Other Indicators

Bowers et al. (2013) provide a comprehensive list of researched early warning indicators, including student and family characteristics, satisfaction with school, grade retention, degree of engagement, and student mobility. We recommend using non-malleable demographic variables like student and family characteristics to compare subgroups of students on the same outcome of interest, rather than relying on these non-malleable demographics to predict who is at risk of dropping out.

Step 1.2 Consider What Types of Indicators and Cut Points Could Apply in Your District

Early warning indicators often vary in their predictive power across different grade levels and districts. The commonly cited early warning indicators can provide a foundation for local decision making, though these indicators will not necessarily be accurate, sufficient, or optimal in all contexts. The primary criterion for determining which indicators to include in an early warning system is the degree to which indicators accurately predict the defined outcome using your own data. To ensure accuracy, different cut points may be needed in different contexts.

Table 1 summarizes commonly cited early warning indicators and cut points for middle and high school grades found to be predictive of high school dropout. The data in your district may lead you to assign a cut point higher or lower than what has been used in the research. Your primary goal when exploring potential early warning indicators should be to weigh what the research says about the indicators and how the indicators could apply in your district.

Table 1. Common Early Warning Indicators and Cut Points for High School Dropout

Indicator	Grade Range	Cut Point
Days absent	6–9	More than 9 per quarter, more than 36 per year[a]
Instructional time	9–12	10% or more[b]
Locally validated behavior data (for example, referrals, suspensions, behavior grades)	9–12	Locally determined[b]
Office referrals	6–9	More than two per quarter, more than six per year[a]
Suspensions	6–9	More than one per quarter, more than two per year[a]
Grade point average	9–12	2.0 or lower on a 4.0 scale[b]
Course failures (any course)	9–12	One or more[b]
Course failures (English language arts and math)	6–8	One or more[a]
Course failure in any core course (English, math, science, and social studies)	9–12	One or more[a,b]
On-Track Indicator	9	Credit deficient for promotion to 10th grade AND one or more failures in core courses[c]

a. Information obtained from Johns Hopkins University (2012).
b. Information obtained from Therriault, O'Cummings, Heppen, Yerhot, and Scala (2013).
c. Information obtained from Allensworth and Easton (2005).

Intended Products

By the end of this section, district staff should have used appendix A to prepare the following:

» A list of indicators you want to explore or refine throughout the other sections of this guide.

Section 2. Select Student-Level Data and Construct Longitudinal Cohorts

What to Expect in Section 2

By the end of this section, you will be able to…

- Select a student cohort based on the outcomes and grade levels of interest (Step 2.1)
- Create a crosswalk of potential indicators and local data sources (Step 2.2)
- Prepare a dataset for analysis (Step 2.3)

Using section 1, you and your colleagues familiarized yourselves with common indicators and determined which types of indicators you want to explore. In section 2, you'll begin to delve into your own data to understand how those indicators can be developed for your unique local setting. This section describes steps that can be taken to understand your own data and prepare data for examining potential early warning indicators. Upon completion of the following steps, you should be able to construct a student-level dataset based on student records that can be used to examine indicators and graduation outcomes.

Steps in this section provide guidance for decisions districts need to make about the specific data used for the analysis and creation of early warning indicators.

Step 2.1 Identify Student Cohorts to Examine

Early warning indicator analyses use historical information from previous student cohorts to develop indicators and then apply the indicators to currently enrolled students or to incoming students. For the purpose of developing or exploring indicators, you need to select one or more prior student cohorts that would have completed the outcome of interest and for whom consistent prior data are available over time.

Select outcome of interest

Common graduation outcomes include on-time graduation, graduation within five or six years, or dropout. The specific outcome you choose will depend on the priorities identified by your district; defining it precisely enough to capture the intended goals of your early warning system will help you develop the most relevant indicators for your local context. It is important that the outcome chosen has a substantial number of students who can achieve it and also a number of students who fail to achieve it.

The selection of outcomes has implications in terms of the number of years of data needed. For example, an analysis that focuses on "on-time graduation" needs at least four years of prior data for the group of students (in order to follow the grade 9 cohort through the year of on-time graduation).

Determine the grade levels of focus

Typically, districts select a group of grade 9 students and build their early warning indicators based on the data collected starting in grade 9. This approach allows districts to identify at-risk students early in high school and without needing to gather data when students are in middle grades. You can choose to select a cohort of middle-grade students if your district aims to intervene with students before they enter high school.

Identify the student cohorts

Data from one cohort of students is the minimum requirement for developing or exploring early warning indicators. We recommend including at least one additional cohort of students so that the indicators developed based on one cohort can be validated with data from another cohort. It's possible you will need to use data from multiple student cohorts for verification purposes.

Districts with small graduation cohorts (for example, smaller than 100) may need to pull data from multiple cohorts to obtain sufficiently large sample sizes. In addition to enrollment, districts need to consider data consistency in deciding the number of cohorts to study and which cohorts to use. It is important to review trends in graduation rates for multiple cohorts to make sure that any anomalous years with particularly high or low graduation rates are not selected to solely represent the local reality.

Step 2.2 Create a Crosswalk of Potential Indicators and Local Data Sources

Creating a crosswalk between your student information system or database and the indicators you identified for exploration will help you understand how many variables and data elements are available for each category of variables, the frequency of data collection, and where these data are located. For example, a crosswalk may reveal that behavior data were collected up until grade 8 or that new disciplinary codes were used starting from a particular school year. The crosswalk should include records relevant to early warning indicators (e.g., attendance, behavior, course performance), and education outcomes as well as student demographics. Table 2 shows how the information gathered from the scan can be documented.

Use the template in appendix B to document the location of the data in your district that will be used to analyze and develop early warning indicators.

Table 2. Example Crosswalk of Data Availability

Variable Type	Variable or Data Element	Years of Collection	Frequency of Data Collection	Source Data File(s) or System	Potential Early Warning Indicator?
Demographics	Race, gender, English language learner status, lunch program status, and students with disabilities	2007–08 to present	Annually	Student Information System (SIS)	No
Attendance	Days present and days missed	2007–08 to present	Daily	SIS	Yes
Behavior	Days suspended and number of suspensions	2007–08 to present	Grading period	Dean of Students	Yes
Course performance	GPA and course grades	2007–08 to present	Grading period	SIS	Yes

Step 2.3 Prepare a Dataset for Analysis

For use in an early warning system, you need a data structure in which all student variables are linked by a unique identifier for each student. Districts that have relational database[2] software programs can create a data structure across multiple years from the annual databases. For districts that have student data dispersed among several different databases, we recommend that you compile these data into one dataset by using computer programs that allow linking data using student identifiers across databases. A common version of a compiled dataset has a single row for each student, with different data elements organized into separate columns. See table 3 for an example.

We generally recommend using data from the entire school year because it offers a more comprehensive and consistent picture of student characteristics, compared to data collected during a single grading period. However, districts should use their best judgment about the unit of time that makes the most sense when compiling data. If the goal is to target students as early as possible, for example, then data that focus on the first grading period may the priority.

2 A system that contains personally identifiable student information, in which all information is linked through a primary key (usually student identification number).

Table 3. Example of Grade 9 Student Data (2011–12) for Students in the Class of 2015

Student ID	Student Name	Days Present	Days Absent	Office Referrals	GPA	Math Grade	Exit Code
3476669	Shawn Lee	175	5	1	3.14	90	Graduated
3476670	Adam Wieters	179	1	0	3.45	82	Graduated
3476671	Sue Grant	142	38	5	1.75	40	Dropped out
3476672	April Brown	160	20	2	2.44	60	Trans-ferred

In preparing data, district staff need to check which variables can be used without further transformation and which variables need to be recoded or constructed. You may need to recode predictor and outcome variables into quantifiable formats. As an example, if a dataset contains grading conventions in the A, B, C, D, and F format, you may want to simply know who has failed a course versus who has not, regardless of the relative grade.

The examples below demonstrate steps needed to recode variables commonly used for early warning indicators:

» Attendance rates may need to be calculated based on days present or days absent.

» Behavior infractions and disciplinary actions are often stored as descriptions or codes. Districts may need to recode the qualitative information into categorical variables, such as the total number of disciplinary actions or whether a student has a disciplinary record.

» Course performance variables vary most across school districts. Depending on the type of data and amount of information available, district staff may be able to convert letter grades into GPA measures, count the number of course failures, count the number of credits accumulated, or create binary variables that indicate course failures.

» The outcome variable for graduation or for other exits needs to be obtained for all students in the sample. Generally, graduation or exit status can be determined from the most recent exit codes. With the final status obtained for all students in the sample, you can decide how to combine or group the final status codes into a few categories such as graduated, dropped out, still actively enrolled, transferred, or unknown. These categories can be further constructed into a final outcome that is binary (1 = achieved the outcome, such as graduated, and 0 = did not achieve the outcome). Exclude data from students whose graduation outcomes cannot be confidently categorized (for example, students who transferred out of the district, students who moved out of the country).

After compiling the dataset and recoding variables, you should conduct a few quality checks to ensure the data are accurate and complete For example, you could compare the percentages of students failing a course against their letter grade to make sure the recording is correct. Combining datasets through merging may be prone to errors, therefore it is important to do quality assurance checks after merging data.

Use the checklist in appendix C to document which data quality checks have been completed.

Intended Products

After completing the three steps described in this section, you should have a database or file that contains the following:

1. Data from one or more longitudinal student cohorts.
2. A unique identifier for each student that is the same for that student across all files.
3. Variables representing outcomes of interest and variables that may predict the outcome.

Section 3. Identify Trends and Patterns Found in Graduation Rates and Indicators

What to Expect in Section 3

By the end of this section, you will be able to…

» Check the distribution of the outcome variable (Step 3.1)

» Examine the distribution of the predictor variables using frequency tables and histograms (Step 3.2)

With the student sample defined and a dataset developed, you can now begin inspecting the trends and patterns for the outcome of interest and potential predictors. This inspection allows you to have a good understanding of your data, which will inform how you choose to develop your predictors into indicators.

Step 3.1 Check the Distribution of the Outcome Variable

For use in an early warning system, it is important to choose an outcome with a fairly large distribution of students who achieve and do not achieve the outcome. If the average dropout rate in your district is low, such as 5 percent, then you may want to choose another outcome, such as on-time graduation, which classifies students into "on-time graduates" and another group for "those who did not graduate on time." The latter group includes both late graduates and non-graduates and thus is larger than the "dropout" group. It is also important to check the distribution of the outcome variable by student subgroups. Once you have reviewed the distribution of the outcome variable, you should have a useful

picture of the percentage of students who fall into the categories with which you have chosen to classify students.

Use appendix D to document the distribution of the outcome variable for one or more cohorts. An additional template for review of the distribution of predictor variables, as described below, is also provided.

Step 3.2 Review the Distribution of the Predictor Variables Using Frequency Tables and Visual Tools

Generally speaking, potential predictors fall into two types: categorical and continuous variables. Categorical variables could include two or multiple categories. Categorical variables that include two categories are considered binary variables. Continuous variables have countless possible values. Variables that have a large number of possible values are usually treated as continuous.

Approaches for reviewing the distribution of predictor variables depends on the types of variables.

- For binary variables, check the prevalence rates—that is, the proportion of students who fall into one category or the other. For example, you can check and document how many students have failed a mathematics course or an English course in grade 9. These prevalence rates can be put in simple frequency tables where the two values and the number and percent of students that exhibit this value are presented in two rows.
- For variables that have a small number of values, such as the number of course failures in a semester, you can make simple frequency tables, in which each value and the number and percent of students that exhibit this value are presented row by row.
- For continuous variables (for example, scores on a test may range from 0 to 100), a frequency table that lists each value is not as useful. You can group several data values into combined groups (see table 4). The range of values within groups can vary if appropriate. For example, if very few students have attendance rates lower than 0.40, then all rates that are below 0.40 can be combined into one group.

You can use the frequency tables to understand distribution, the central tendency, and the dispersion of the variables. First, the number or percent of students who had each value or were within each value range tells the distribution of the variable. Second, the frequency table can reveal the "center" of a distribution, or the most frequently occurring value or value range. Third, the highest and lowest values show the spread of the data.

Table 4. Sample Frequency Table of Students by Exam Score Ranges

Exam Score	Frequency	Percentage of Students
90–100	45	12%
80–89	80	22%
70–79	125	34%
60–69	70	19%
50–59	30	8%
0–49	20	5%

After summarizing the data into categories and documenting the frequencies and percentages (table 4), you can create bar charts. Bar charts visually depict the number or percent of students within categories of variable. For example, figure 1 shows that grade 9 absences were spread widely from 0 to 4 days to more than 40 days. Figure 2 shows a narrower range of absences among students, with a majority of students absent in the range of 5–9 days.

Bar charts can be created with spreadsheet and statistical software. You can use bar charts for different cohorts or subgroups of students to show how frequencies vary across groups. With most statistical software tools, you can also create histograms to summarize the shape and spread of the distributions of continuous quantitative data without breaking the data into categories first.

Bar charts and histograms are also useful tools for understanding the values or cut points that indicate risk among your students, since these charts can show how many students will be "flagged" as at risk by choosing a particular cut point. For example, if you choose to use 4 percent as the cut point to indicate frequent absences, then 22 percent of students in figure 1 will be classified as "not frequently absent" while the rest will be classified into the "frequently absent" group.

As you examine your data, you can use information revealed by frequency tables, bar charts, and/or histograms to make decisions about variable selection or recoding. For example, you may choose to collapse some smaller value ranges into a larger group. Instead of having 50–59 and 0–49 as two separate value ranges (as in table 4), you may choose to combine them into one group. With a distribution of absence rates depicted in figure 1, you may choose to group students with 0–4 and 5–9 absences as "not frequently absent," and combine the rest as "frequently absent." You can also choose to drop some variables that have little variation. For example, if only 3 out of 2,000 students were expelled, this variable is unlikely to be useful.

Inspecting the distribution of variables using frequency tables and visual tools will help you to understand the nature of the variables in your data and further explore the relationships between predictor variables and the outcome of interest.

Figure 1. Example Bar Chart of Grade 9 Absences, Scenario 1

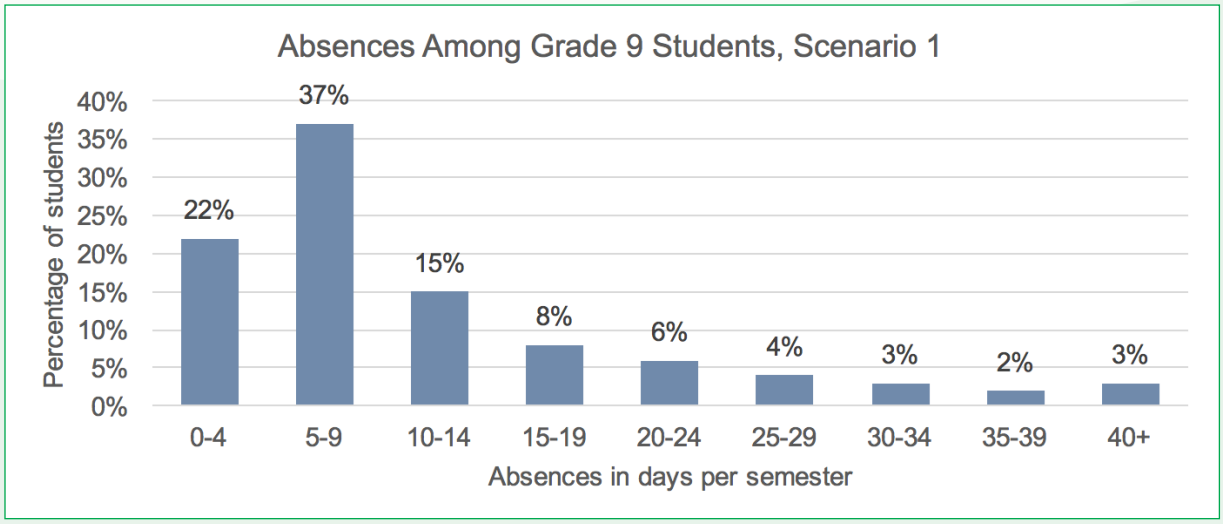
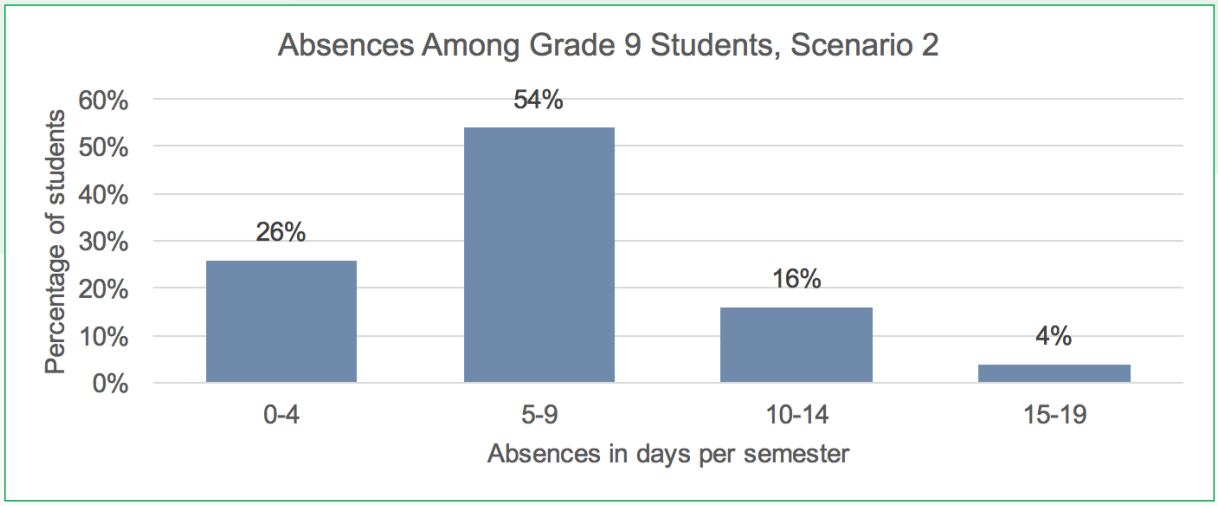


Figure 2. Example Bar Chart of Grade 9 Absences, Scenario 2



Use the bar chart feature of the spreadsheet software to create graphs for each early warning indicator being analyzed.

Intended Products

After completing the two steps described in this section, you should have the following:

1. A spreadsheet with frequency tables for each variable under consideration to be used to develop early warning indicators in the next section.

2. Bar charts depicting the distribution of student-level data associated with each potential early warning indicator.

Section 4. Analyze Relationships Between Potential Predictors and Graduation Outcomes

What to Expect in Section 4

By the end of this section, you will be able to…

- Summarize potential predictors for students who achieved the graduation outcome and those who did not (Step 4.1)
- Visually depict relationships between potential predictors and the outcome (Step 4.2)
- Identify predictors with meaningful differences across the two outcome groups (Step 4.3)

In this section, you will use appendix E to work through the summary, visual depiction, and comparison of your potential predictors and your outcome groups.

Step 4.1 Summarize Potential Predictors for Students Who Achieved the Graduation Outcome and Those Who Did Not

For continuous variables such as attendance rates, numbered course grades, and test scores, you can calculate the means for students who achieved the graduation outcome and means for students who did not. A sample table is presented in table 5. The table suggests that attendance, GPA, and number of course failures are related to the graduation outcome. The table also suggests it would be difficult to predict the graduation outcome for a student based on the number of suspensions he or she receives.[3]

For categorical predictor variables, creating simple contingency tables will summarize the relationships between the predictor and the outcome. A sample contingency table that summarizes the relationships between a binary predictor and the outcome is presented in table 6. The table shows that among students who failed one or more courses in grade 9, only 15 percent graduated on time, whereas 85 percent of students who did not fail any course graduated in the same time frame. Among those who did not graduate, 62 percent had failed one or more courses in grade 9. Additionally, the table shows that about two thirds of students graduated on time (1,000 out of 1,500) and 30 percent of students failed at least one course in grade 9 (460 out of 1,500). These numbers suggest that students with course failures were much less likely to graduate on time from high school than students without course failures (15 percent vs. 85 percent), which, in turn, suggests that course failure is related to graduation and thus could be an accurate indicator.[4]

Table 5. Average Scores of Predictors by Graduation Status

Predictor Variables	Students Achieving the Graduation Outcome	Students Not Achieving the Graduation Outcome
Attendance rate	98%	75%
Grade 9 GPA	2.8	1.7
Number of course failures	0.3	2.7
Number of suspensions	0.9	1.0

Use appendix E to document graduation outcomes of binary predictors in contingency tables.

3 The difference in the means between groups can be further examined using statistical tests. T-tests, for example, test whether the difference in means between two groups is statistically significant.

4 The strength of the relationships depicted in the table can be further examined using statistical tests. Chi-square tests, for example, examine whether the relationships between categorical variables are statistically significant.

Table 6. Graduation Outcomes by Course Failure Status

Examples of Binary Predictors	Graduation Outcomes-Graduated On Time Number	Graduation Outcomes-Graduated On Time Percentage	Graduation Outcomes-Did Not Graduate Number	Graduation Outcomes-Did Not Graduate Percentage	Total
Failed one or more courses in grade 9	150	15%	310	62%	460
Did not fail any course	850	85%	190	39%	1,040
Total	1,000	100%	500	100%	1,500

Step 4.2 Visually Depict Relationships Between Potential Predictors and the Outcome

Bar charts can be used to visually display relationships between potential predictors and outcomes. Continuous predictor variables can be recoded so values of the variable fall into groups for creation of bar charts. You can use the visual depiction from a bar chart to understand the nature and strength of the relationship. Table 7 and figures 3 and 4 illustrate two situations of relationships between absence rates and graduation rates. If the data show a pattern like that described in situation 1 (figure 3), then you might conclude that the lower a student's absence rate, the more likely he or she will graduate from high school on time. However, if the data show a pattern that is more similar to situation 2 (figure 4), then the absence rate does not seem related to graduation status. Categories used to create the bars in these figures are hypothetical examples. You can use any categories that are appropriate in your context to create your own bar charts.

Table 7. Examples of Relationships Between Absence and Graduation Rates

Attendance Rate	Graduation Rate: Situation 1	Graduation Rate: Situation 2
0% to 4%	90%	78%
5% to 9%	81%	81%
10% to 14%	70%	75%
15% to 19%	55%	77%
20% or more	40%	74%

Figure 3. Bar Chart of Graduation Rates by Attendance Rates, Situation 1

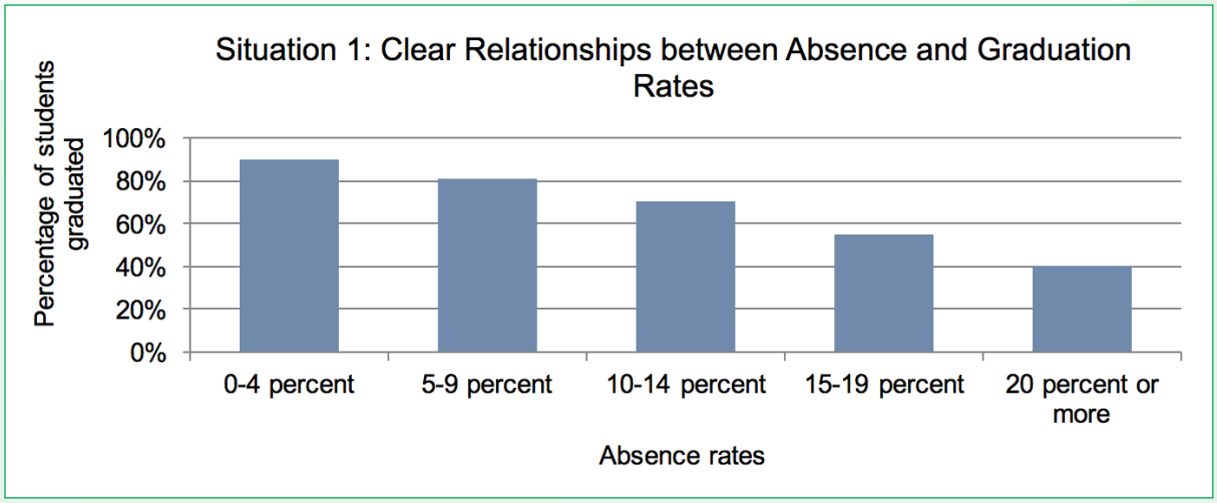
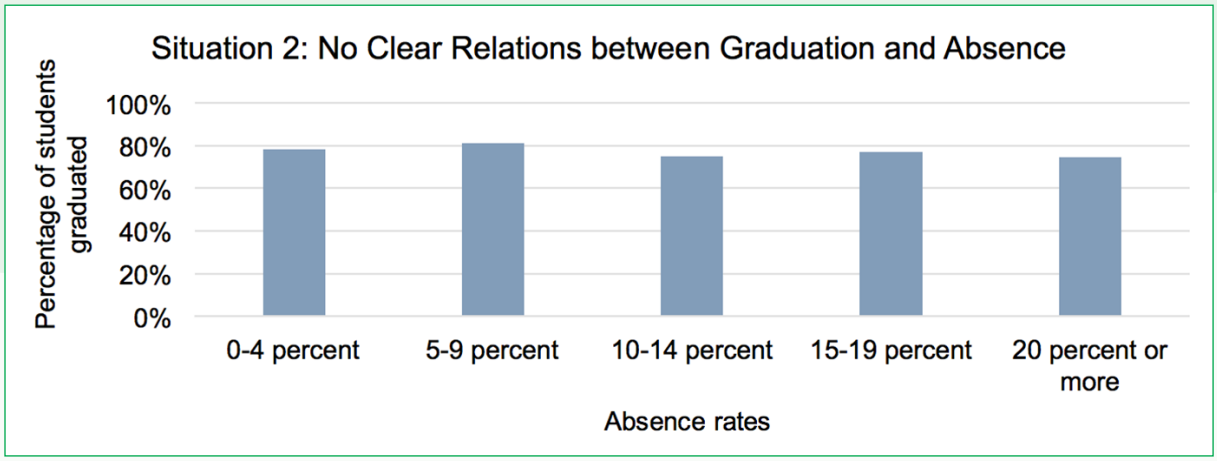


Figure 4. Bar Chart of Graduation Rates by Attendance Rates, Situation 2



Step 4.3 Identify Predictors with Meaningful Differences Across the Two Outcome Categories

The final step in this section involves interpreting the results from section 3 and steps 4.1 and 4.2. In order for an early warning system to function usefully, it must be based on predictor variables that show large or meaningful differences in whether students achieved the graduation outcome or not. There is not a specific standard on what counts as a "meaningful difference," so you will need to make judgments based on the information examined. In the example provided in table 5, there were large differences among students who achieved the graduation outcome and those who did not, specifically on the variables of attendance rate, GPA, and course failures. These likely would represent meaningful differences. The following sections provide guidance on considering "how large is large enough." In preparation for that work, take some time to note where large differences appear in your tables and graphs, and discuss them with colleagues.

Intended Products

After completing the three steps in this section, you should have the following:

1. Tables of predictors that show differences between the group that achieved the outcome and the group that did not achieve the same outcome.

2. Histograms displaying the relationship between different indicators and graduation outcomes.

3. Lists of variables with meaningful differences between students who did or did not meet graduation outcomes.

Section 5. Choose the Best Cut Point for Early Warning Indicators

What to Expect in Section 5

By the end of this section, you will be able to…

- Create indicators based on cut points of predictors (Step 5.1)
- Check the performance of the indicators against the actual outcome (Step 5.2)
- Compare classification performance based on common measures (Step 5.3)

Step 5.1 Create Indicators Based on Cut Points of Predictors

The identified predictors can be grouped into binary variables, discrete variables with more than two but small number of values, and continuous variables. Some of these have built-in hierarchies and thus make assigning the cut point straightforward (for example, "Yes or No" on a course failure variable already tells you where students fall on that measure). For other variables that contain a range of values, a good starting point would be to consult table 1 of this guide, "Common Early Warning Indicators and Cut Points for High School Dropout." You may also decide to try cut points based on the distribution of your own data. Multiple binary variables could be created using values of a single discrete or continuous variable.

The value used to create a binary variable is a cut point. The goal of this step is to select the cut point that creates the most appropriate indicator. Assigning the cut points is the final step in turning your variables into early warning indicators. Below are some examples for how you can do this.

- Binary predictors can be used as indicators without further construction because they are already in a 1/0, or yes/no, format.
- Discrete variables have a limited number of data values, such as the number of course failures or number of disciplinary actions. You can use these values to create multiple indicators based on the same variable. Each value within the variable can be used to create an indicator. The indicators based on the same variables then can be compared against common performance criteria specified in the next section.
- Continuous variables with a whole set of values can either identify the optimal cut point using Receiver Operating Characteristic (ROC) curve analysis or select values based on the distribution of the variables (section 3) to create indicators.

Depending on the statistical capacity of your district, you can identify optimal cut points for continuous variables through ROC curve analysis or manual calculation (an introduction of ROC analysis is included in appendix F). If the statistical capacity of your district is more limited, then you could choose several cut points based on distribution of your data and create an indicator for each of the selected cut points. For example, you may designate 95 percent, 90 percent, 85 percent, 80 percent, 75 percent, or 70 percent as potential cut points for attendance rates, and then choose one that yields the best results based on the common performance criteria described in steps 5.2 and 5.3.

Step 5.2 Check the Performance of the Indicators Against the Actual Outcome

After creating the indicators, you should apply the indicators to the same dataset and the same cohort of students identified in step 2, and compare the predictions against their actual outcomes of interest to evaluate the quality of the classifications made by the indicators. Each indicator classifies each student as "likely to achieve the outcome" or "unlikely to achieve the outcome." This classification may or may not match the student's actual outcome. Table 8 demonstrates the four aspects of classification quality you will need to examine using "dropping out of school" as an example outcome. The goal is to maximize correct classifications (true positive and negative) and minimize incorrect classifications (false positive and negative).

Table 8. Event Table for Calculating Dropout Contingency Proportions

Indicator Prediction	Actual Outcome Dropped Out	Actual Outcome Did Not Drop Out	Total
Likely to dropout	True positive (TP)	False positive (FP)	TP+FP
Unlikely to dropout	False negative (FN)	True negative (TN)	FN+TN
Total	TP+FN	FP+TN	TP+FP+FN+TN

Note. TP = True positive: Dropouts correctly identified as at risk of dropping out. FN = False negative: Graduates incorrectly identified as at risk of dropping out. TN = True negative: Graduates correctly identified as not at risk of dropping out. FP = False positive: Dropouts incorrectly identified as not at risk of dropping out.

Table 9 indicates the results of the classification using the four categories described above (table 8). Once you compile your data in a similar table you can count the number of true negative, false negative, false positive, and true positive results. Counting these values will give you a sense of how accurately the selected indicator performed in predicting the actual outcome.

Table 9. Example of Students Flagged by an Indicator Compared Against Actual Graduation Outcome

Student	At Risk - Indicator: Failed a Core Course	Actual Outcome - Dropped Out	Results of Classification
Student A	0	0	True negative
Student B	1	1	True positive
Student C	0	1	False negative
Student D	0	0	True negative
Student E	0	0	True negative
Student F	1	1	True positive
Student G	0	0	True negative
Student H	1	1	True positive
Student I	1	0	False positive
Student J	0	1	True negative
Student K	0	0	True negative
Student L	0	0	True negative
Student M	0	0	True negative
Student N	1	0	False positive

Note. There are three TP, eight TN, one FN, and two FP.

Step 5.3 Compare Classification Performance Based on Common Measures

Tallying the classification results is a useful way to summarize indicator performance, but you may be wondering what to do with this information. No early warning indicator will correctly predict 100 percent of outcomes. In order to identify optimal cut points, you need to calculate sensitivity, specificity, and Youden index or other global measures based on the number of true negative, false negative, false positive, and true positive results.

- Sensitivity measures the proportion of positives that are correctly identified as such, in this case, the percentage of dropouts who are correctly identified as at risk of dropping out. The equation to calculate sensitivity is TP/(TP + FN).
- Specificity measures the proportion of negatives that are correctly identified as such, in this case, the percentage of graduates who are correctly identified as not at risk of dropping out. The equation to calculate specificity is TN/(TN + FP).
- Youden index or other global measures take into account sensitivity and specificity. W. J. Youden (1950) suggested using "sensitivity + specificity − 1" to measure the performance of binary indicators. This measure has been named the "Youden index." The closer the value on the Youden index is to 1, the more useful the indicator. A value of 1 suggests that there are no false positive

or false negative results. We suggest using this index instead of sensitivity or specificity alone when choosing optimal cut points. The cut point that results in the largest value on the Youden index is the optimal cut point for that particular indicator.

An example summary of number and percentage of grade 9 students who graduated above and below GPA cut points and the performance of the indicators based on these cut points is provided in table 10. Each cut point can be used to create a GPA indicator. For example, the cut point of 2.3 can create an indicator of "low GPA," with students receiving a GPA of 2.3 or lower coded as "low GPA" and students receiving a GPA higher than 2.3 coded as not having "low GPA." If the cut point of 2.3 is used, the indicator can correctly identify 44 percent of dropouts as at risk and 93 percent of non-dropouts as not at risk. This cut point is optimal because it best balances sensitivity and specificity by having the highest value on the Youden index. A template for calculating these measures is in appendix F.

Intended Products

After completing the three steps in this section, you should have the following:

1. A list of indicators and their performance results in terms of sensitivity, specificity, and Youden index values.

2. The cut points that are used to create the indicators that are based on discrete or continuous variables.

Table 10. Numbers and Percentages of Grade 9 Students Who Dropped Out Above and Below GPA Thresholds (Pseudo Data) and Scores on Measures of Performance of Indicator Based on the Listed GPA Cut Points

Variable GPA	Dropouts - Students Who Displayed Indicator Percentage (N)	Students Who Did Not Display Indicator Percentage (N)	Graduates - Students Who Displayed Indicator Percentage (N)	Graduates - Students Who Did Not Display Indicator Percentage (N)	Measures of Performance of Indicator Based on GPA Cut Points - Sensitivity	Measures of Performance of Indicator Based on GPA Cut Points - Specificity	Measures of Performance of Indicator Based on GPA Cut Points - Youden Index
1.5	41% (87)	5% (527)	59% (125)	95% (10,006)	0.14	0.99	0.13
1.6	39% (100)	4% (420	61% (157)	96% (10,068)	0.19	0.98	0.17
1.7	37% (120)	4% (417)	63% (204)	96% (10,004)	0.22	0.98	0.2
1.8	33% (137)	4% (413)	67% (279)	96% (9,916)	0.25	0.97	0.22
1.9	31% (157)	4% (410)	69% (348)	96% (9,830)	0.28	0.97	0.25
2	29% (174)	4% (406)	71% (427)	96% (9,738)	0.30	0.96	0.26
2.1	26% (187)	4% (401)	74% (533)	96% (9,624)	0.32	0.95	0.27
2.2	25% (206	4% (397)	75% (620)	96% (9,522)	0.34	0.94	0.28
2.3	24% (231)	3% (293)	76% (733)	97% (9,488)	0.44	0.93	0.37
	True positive (accurate identification)	False negative	False positive	True negative (accurate identification)			

Note. GPA of 2.3 in the last row (highlighted in orange) has the best score on sensitivity and Youden Index.

Section 6. Test the Relative Performance of Early Warning Indicators

What to Expect in Section 6

By the end of this section, you will be able to…

- Choose a criterion that is aligned with the local priorities (Step 6.1)
- Compare the indicators and actual student outcome (Step 6.2)
- Decide the number of indicators or combination of indicators (Step 6.3)
- Decide the level of risk on which the system should focus (Step 6.4)

Step 6.1 Choose a Criterion That Is Aligned With the Local Priorities

Similar to the performance-based comparisons done to select cut points within indicators, the comparisons between indicators can be based on the same criteria, but the criteria should be aligned with the local priorities in addition to performance. Step 5.3 describes how to calculate sensitivity, specificity, and Youden index for each indicator. After summarizing indicators' performance on these criteria, you need to decide which criterion or criteria to focus on. On the one hand, you may want to select a criterion that allows you to identify indicators that flag only students who are most at risk of not achieving the outcome of interest so that resources are not focused on students who would have graduated without additional intervention. On the other hand, you also need indicators to be able to identify a large enough proportion of students who cannot achieve the outcome so that students who need support are able to receive the service. We recommend using a global measure such

as the Youden index that balances the two aspects of needs. However, the choice of criterion should also take into account the goals of the district.

Before you dive into this work, you should consider your local priorities.

- If the priority is to target only students who are at the highest risk of not achieving a graduation outcome, then an indicator that is highly sensitive or identifies the smallest proportion of students as at risk is the best choice.
- If the priority is to identify and provide services to all grade 9 students who are at risk of not achieving a graduation outcome, then it is more appropriate to select an indicator that is highly specific or an indicator that identifies a large number of students who are at risk.
- If the priority is to balance both targeting students at the highest risk of not achieving a graduation outcome and identifying and providing services to all grade 9 students at risk of not achieving a graduation outcome, then an indicator that is high on both sensitivity and specificity is most appropriate.
- If the priority is to best align student needs with intervention then it is more appropriate to select indicators based on the fit between the indicators and the interventions available. For example, if the current intervention capacity is concentrated on behavior, district staff may choose to focus on behavior indicators such as attendance and disciplinary actions. If the current intervention capacity is centered on academic support, then indicators based on course failure or performance can identify students who need academic support.

Use the decision matrix in appendix G to match priorities for your indicator to an appropriate indicator criterion.

Once the criterion or criteria are chosen, you can compare the indicators against each other. For example, if your district would like to target a small number of students to provide service, an indicator that flags 5 percent of students is more appropriate than an indicator that flags half of the cohort. Or if your district plans to provide one-on-one tutoring to students who failed algebra, then the "attendance" indicator or "low GPA" indicator are not as good as the indicator of "failing a math course."

In addition to the individual indicators, combinations of indicators can be compared against each other based on sensitivity, specificity, and Youden index. For example, you can compare the performance of "low GPA, low attendance, and failed a core course" with "low GPA, failed a math course, and was suspended more than once" by counting the number of students who were flagged by these two sets of indicators and calculating sensitivity, specificity, and Youden index (following the steps specified in section 5). The set of indicators that have relatively better performance should be selected.

The template for calculating performance measures of indicators based on different cut points (table F1 in appendix F) can be adopted for this purpose in order to compare and rank the indicators. You should rank the indicators based on the chosen performance measure you identified for each one in section 5. Table G1 (see appendix G) provides a template for comparing and ranking the indicators based on the chosen measure. Unlike table F1, which documents the performance measures of the indicators, Table G1 lets you prioritize indicators that perform well in the most important criterion or criteria. You can work through this template to see which combination of indicators best predicts the outcome you've chosen to focus on.

Step 6.2 Compare the Indicators and Actual Student Outcome

After the indicators are ranked, the next step is to count the percentage of students who are flagged by each indicator. You can document the ranking of indicators in terms of performance and the percentage of students they flag, and then make a decision that balances indicator performance and practicality. Indicators, even though identified as the best performing in terms of sensitivity and specificity, may flag too many or too few students who actually failed a graduation outcome. This is where local discretion comes in—you should set a target proportion of students, based on your current capacity that you can realistically direct toward an appropriate intervention given resource restraints. An indicator that flags almost everybody does not work well. Similarly, if an indicator only flags a handful of students, then it may not be as helpful. Thus, you may select an indicator that does not have the largest value on the Youden index but nonetheless identifies a practical number of students for intervention that your district's current capacity can accommodate.

In addition, you should cross-tabulate students flagged by different indicators counting which students are identified by each indicator. If multiple indicators flag the exact same group of students, then only one indicator is needed. This tends to happen when two indicators are within the same category (e.g., the indicator of "failed a core course" and "failed two or more courses" are both based on course performance variables). It is recommended not to include two highly similar indicators in the system because adding both does not provide more accuracy.

Step 6.3 Decide the Number of Indicators or the Combination of Indicators

Research suggest that three or four indicators are sufficient to form a robust early warning system.[5] You can decide on the number of indicators based on the size of the group of students flagged by each indicator. You also need to decide on which indicators to choose among the ones that do a good job predicting outcomes identified in section 5. When indicators are similar on the performance criterion, you should try different combinations of indicators and choose a set that identifies at-risk students efficiently and effectively based on the priorities your district has identified.

Step 6.4 Decide the Level of Risk on Which the System Should Focus

After choosing a set of indicators, district staff can compare the graduation outcomes for students with different numbers of flags in two ways. First, you can count the number and percentage of students with different numbers of flags who achieved the desired graduation outcome and who did not (see table 10 in section 5). Second, you can count how many dropouts received no flag, one flag, or more than one flag.

Table 11 shows the percentage of students with different numbers of indicators, comparing the rates for dropouts and those who graduated. You should try different combinations, selecting several from the possible combinations and replicating table 11, to gauge the indicators overall utility as a set. Comparing different combinations of indicators will help determine which set would work best and would meet your local priorities.

5 Jobs for the Future. (2014). *Early warning indicators and segmentation analysis: A technical guide on data studies that inform dropout prevention and recovery.* Retrieved on July 25, 2016 from https://www2.ed.gov/programs/dropout/earlywarningindicators.pdf

Based on the distribution of numbers of flags, you can then decide whether you should target the students with multiple flags or target students who received any flags. Table 12 shows an example of how you could document this work. The table also shows that combinations of flags seem to capture a larger proportion of students who did not achieve a graduation outcome.

Table 11. Graduation Outcomes for Students With Different Numbers of Flags on Four Selected Indicators (maximum number of flags is four)

Number of Flags	Percentage of Students Who Dropped Out	Percentage of Graduates
No flag	39%	61%
One flag	61%	39%
Two flags	79%	21%
Three flags	91%	9%
Four flags	99%	1%

Table 12. Graduation Outcomes for Students With Different Combinations of Flags on Three Selected Indicators

Combination of Flags	Percentage of Students Who Dropped Out	Percentage of Graduates
No flag	15%	85%
Flag A	27%	73%
Flag B	22%	78%
Flag C	31%	69%
Flags A & B	35%	65%
Flags A & C	46%	54%
Flags B & C	79%	21%
Flags A, B, & C	91%	9%

Note. Flag A: attendance below 90 percent, flag B: being suspended once or more, flag C: failed a core course.

After completing these tables, you should have a sense of how the early warning indicators that have been developed would function for flagging at-risk students. This information will be important as you test how these indicators work in other grade levels, with other student cohorts or subgroups of students, and for other graduation outcomes using your own data in section 7.

Intended Products

After completing the four steps in this section, districts should have the following:

1. Three or four early warning indicators.
2. A distribution of students with different numbers of flags.

Section 7. Examine the Reliability and Accuracy of Different Indicators or Combinations of Indicators

What to Expect in Section 7

By the end of this section, you will be able to…

- Apply the indicators to data collected at different grade levels from the same students (Step 7.1)
- Apply the indicators to a different cohort of students (Step 7.2)
- Refine and reflect on indicators (Step 7.3)

Step 7.1 Apply the Indicators to Data Collected at Different Grade Levels from the Same Students

District staff can check the reliability of indicators by applying the same indicators to data collected at other grade level(s) for the same cohort of students. For example, the same set of grade 9 indicators developed for the class of 2015 can be created based on the grade 10 data of the same cohort. If the indicators are perfectly reliable and students' performance is stable, students who are flagged as at risk in grade 9 should be flagged as at risk by the same indicator, such as GPA below 2.3, based on their performance in grade 10 or 11. It is possible that some students become at risk by falling further behind or catch up in later grades and no longer are at risk; however, most of the students (for example, 60 percent to 70 percent or more) typically maintain the same risk status without intervention in place. If the proportions of students being identified using data collected from later or earlier grade levels are similar to the percentages of students being identified using the current grade level, then one can conclude that the indicators are reliable.

Step 7.2 Apply the Indicators to a Different Cohort of Students

District staff can verify the accuracy of indicators by applying the same indicators to data collected from a different cohort of students. If the indicators can establish similar levels of accuracy when applied to another cohort, you can confidently apply the indicators to current students.

Step 7.3 Refine and Reflect on Indicators

Finally, an early warning indicator is only as good as the data it is built on. With this in mind, you should revisit the analyses in a few years (we recommend at most 5 years) to check whether the cut points need to be adjusted as the student population shifts. Also, the reliability of indicators needs to be checked by student subgroups. For example, does your indicator for "low GPA" (defined as a GPA below 2.3, based on your indicator construction) flag a similar number of students across all subgroups? Does the "low GPA" flag disproportionally large numbers of students who are English learners or students with special needs? We encourage districts to reflect on whether the accuracy of the early warning indicators is high enough to justify its implementation and to justify its implementation across student subgroups. There are few, if any, "right" answers or obvious solutions to these guiding questions. Therefore, the trade-offs of these decisions must be carefully considered.

Intended Products

After completing the three steps in this section, districts should have the following:

» A set of early warning indicators that are reliable and accurate.

Conclusion

Establishing a set of early warning indicators based on local context can be a powerful tool for alerting educators to students who need intervention to graduate on time before it is too late. This guide leads school districts and their staff through the technical steps involved in identifying early warning indicators based on the local context that can power a robust and accurate early warning system. We hope that this guide provides useful information and templates for districts to develop their own early warning indicators. It is important to note that identifying at-risk students is only the start to developing an early warning system. After students have been identified, they must receive appropriate support and monitoring to get back on track for high school graduation.

References

Alexander, K. L., Entwisle, D. R., & Kabbani, N. S. (2001). The dropout process in life course perspective: Early risk factors at home and school. *Teachers College Record, 103*, 760–823.

Allensworth, E. M., & Easton, J. Q. (2005). *The on-track indicator as a predicator of high school graduation*. Chicago, IL: University of Chicago Consortium on Chicago School Research. Retrieved from https://consortium.uchicago.edu/sites/default/files/publications/p78.pdf

Allensworth, E. M., & Easton, J. Q. (2007). *What matters for staying on-track and graduating in Chicago public high schools: A close look at course grades, failures, and attendance in the freshman year.* Chicago, IL: University of Chicago Consortium on Chicago School Research. Retrieved from http://ccsr.uchicago.edu/sites/default/files/publications/07%20What%20Matters%20Final.pdf

Balfanz, R., Herzog, L., & Mac Iver, D. J. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educational Psychologist, 42*(4), 223–235.

Battin-Pearson, S., Newcomb, M. D., Abbott, R. D., Hill, K. G., Catalano, R. F., & Hawkins, J. D. (2000). Predictors of early high school dropout: A test of five theories. *Journal of Educational Psychology, 92, 568*–582.

Bowers, A. J., Sprott, R., & Taff, S. A. (2013). Do we know who will drop out? A review of the predictors of dropping out of high school: precision, sensitivity, and specificity. *The High School Journal, 96*(2), 77–100.

Carl, B., Richardson, J. T., Cheng, E., Kim, H., & Meyer, R. H. (2013). Theory and application of early warning systems for high school and beyond. *Journal of Education for Students Placed at Risk (JESPAR), 18*(1), 29–49.

Eng, J. (2005). Receiver Operating Characteristic analysis: A primer. *Academic Radiology, 12*, 909–916.

Fedorowicz, A. E. (1995). *Children's coping questionnaire (CCQ): Development and factor structure* (Doctoral dissertation, Simon Fraser University, Buranaby, British Columbia).

Fleming, C. B., Haggerty, K. P., Catalano, R. F., Harachi, T. W., Mazza, J. J., & Gruman, D. H. (2005). Do social and behavioral characteristics targeted by preventive interventions predict standardized test scores and grades? *Journal of School Health, 75*(9), 342–349.

Frazelle, S., Nagel, A., & Northwest, E. (2015). *A practitioner's guide to implementing early warning systems* (REL 2015–056). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest.

Hartman, J., Wilkins, C., Gregory, L., Gould, L. F., & D'Souza, S. (2011). *Applying an on-track indicator for high school graduation: adapting the Consortium on Chicago School Research indicator for five Texas districts* (Issues & Answers Report, REL 2011–No. 100). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from http://ies.ed.gov/ncee/edlabs

Heppen, J. B., & Therriault, S. B. (2008). *Developing early warning systems to identify potential high school dropouts*. Washington, DC: National High School Center. Retrieved from http://files.eric.ed.gov/fulltext/ED521558.pdf

Jobs for the Future. (2014). *Early warning indicators and segmentation analysis: a technical guide on data studies that inform dropout prevention and recovery.* Retrieved on July 25, 2016 from https://www2.ed.gov/programs/dropout/earlywarningindicators.pdf

Johns Hopkins University, School of Education, Center for Social Organization of Schools. (2012). *Using data to keep all students on track to graduation: Team playbook.* Baltimore, MD: Author. Retrieved from http://new.every1graduates.org/wp-content/uploads/2012/11/Team_Playbook.pdf

Knowles, J. E. (2015). Of needles and haystacks: Building an accurate statewide dropout early warning system in Wisconsin. *Journal of Educational Data Mining,* 7(3), 18–67.

Neild, R., & Balfanz, R. (2006). *Unfulfilled promise: The dimensions and characteristics of Philadelphia's dropout crisis, 2000–2005.* Baltimore, MD: Center for Social Organization of Schools, Johns Hopkins University. Retrieved from http://files.eric.ed.gov/fulltext/ED538341.pdf

Norbury, H., Wong, M., Wan, M., Reese, K., Dhillon, S., & Gerdeman, R. (2012). *Using the freshman on-track indicator to predict graduation in two urban districts in the Midwest Region* (Issues & Answers Report, REL 2012–No. 134). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest.

Roderick, M. (1993). *The path to dropping out: Evidence for intervention.* Westport, CT: Auburn House.

Stuit, D., O'Cummings, M., Norbury, H., Heppen, J., Dhillon, S., Lindsay, J., & Zhu, B. (2016). *Identifying early warning indicators in three Ohio school districts* (REL 2016–118). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved from http://ies.ed.gov/ncee/edlabs/

Therriault, S. B., O'Cummings, M., Heppen, J., Yerhot, L., & Scala, J. (2013). *High school early warning intervention monitoring system implementation guide.* Washington, DC: American Institutes for Research, National High School Center. Retrieved from http://files.eric.ed.gov/fulltext/ED521686.pdf

Uekawa, K., Merola, S., Fernandez, F., & Porowski, A. (2010). *Creating an early warning system: Predictors of dropout in Delaware* (REL Mid-Atlantic Technical Assistance Brief, REL MA 1.1.75-10). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic.

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer,* 3, 32–35. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/1097-0142(1950)3:1%3C32::AID-CNCR2820030106%3E3.0.CO;2-3/pdf

Appendices

The appendices are based on the outcome of on-time graduation. An editable version of each planning form in these appendices is available in the companion document, *District Guide for Creating Indicators for Early Warning Systems — Editable Planning Forms*.

Appendix A. Early Warning Indicator Documentation Template (from Section 1)

Drawing on information shared in section 1, use this template (table A1) to organize the early warning indicators used in the research into "ABC" categories and identify the variables your district is interested in developing into early warning indicators. District staff can group the indicators into broad categories, including attendance, behavior, course performance, and composite/other. When documenting the early warning indicators, please record the following:

» **Grade range:** An individual grade or range of grades (for example, grades 9–12). Indicators may have different cut points at different grade ranges. In this case, list indicators with unique grade range and cut point combinations on separate rows.

» **Time frame:** The time(s) during the school year at which an existing indicator was applied may differ by indicator. Some indicators might have been collected every grading period, whereas others may only be available at the end of a year. In many instances, the indicators may become available on multiple occasions.

» **Cut point:** The threshold, based on research, at which students are flagged in early warning indicators for being at risk. Be sure to indicate the polarity of the cut point (above or below).

Table A1. Template for Documenting Indicators Used in the Research

Indicator	Grade Range	Time Frame	Cut Point	Interested in Exploring?
Attendance				
Behavior				
Course Performance				
Composite Indicators/ Others				

Note. An editable version of this table is available in *District Guide for Creating Indicators for Early Warning Systems — Editable Planning Forms*.

Appendix B. Map of Data Availability (from Section 2)

Use this template (table B1) to identify the specific data elements comprising the indicators and variables that will be used to create the early warning indicators. For each data element, note the number of years the data have been collected and specific location of the files. Use separate rows as necessary to list the data elements that are located in different files or locations.

Table B1. Template: Data Elements, Data Collection Schedules, and Locations

Indicator or Variable	Data Elements	Frequency of Collection	Location of Files
Demographics			
Attendance			
Behavior			
Course Performance			
Composite Indicators/Other			

Note. Demographic variables will not be used to create indicators. We recommend gathering them so they can inspect patterns in outcome and predictors between student subgroups based on demographic characteristics (such as gender and racial group).

An editable version of this table is available in *District Guide for Creating Indicators for Early Warning Systems — Editable Planning Forms*.

Appendix C. Data Quality Checklist (from Section 2)

Step 2.3 of the guide describes the process of conducting data quality checks. Once the requisite data to be analyzed have been merged into a single dataset, conduct a few quality checks to ensure accuracy and availability. Quality checks will identify missing data for students, for particular grading periods and school years, and for all early warning indicators and outcome variables. Select the checkbox next to each quality check when the process has been completed.

- ☐ **All student data being analyzed for all selected years and grading periods are available (based on appendix B: Map of Data Availability).**
- ☐ **All student records have a matching Student ID.**
- ☐ **All variables use a consistent format (for example, all dates use the MM/DD/YYYY format).**
- ☐ **All student records include all variables being analyzed; the student records do not include gaps in any variable being analyzed.**
- ☐ **Student records are continuously present across grading periods and school years.**

Note. An editable version of this checklist is available in *District Guide for Creating Indicators for Early Warning Systems — Editable Planning Forms*.

Appendix D. Documenting the Distribution of the Outcome Variable (from Section 3)

Use this template (table D1) to record the distribution of the outcome variable. Replace the column headings with any outcome variable, such as four- or five-year graduation rates, that your district has selected based on the guidance from section 3.

Table D1. Distribution of Outcome Variable: High School Graduation and Dropout Rates

Cohort	Graduation Rate	Four-Year Dropout Rate	Five-Year Dropout Rate
Example: Class of 2015	*97.5%*	*3.3%*	*2.7%*
Class of 20_____			
Class of 20_____			
Class of 20_____			
Class of 20_____			
Average			

Note. An editable version of this table is available in *District Guide for Creating Indicators for Early Warning Systems — Editable Planning Forms*.

Use the frequency table (table D2) to display the distribution of predictor variables (step 3.2) and create a line or bar chart with the percentage of students grouped by the predictor variable ranges you have selected (step 3.3).

Table D2. Frequency Table of Students by Predictor Variable

Predictor Variable Range	Frequency	Percentage of Students
Example: 0–4 days absent	*375*	*47.8%*

Note. An editable version of this table is available in *District Guide for Creating Indicators for Early Warning Systems — Editable Planning Forms*.

Appendix E. Graduation Outcomes Tracker (from Section 4)

Use table E1 to document the relationship between the selected binary predictors and graduation outcomes. Replicate the table for each indicator you will explore for a relationship to graduation or dropout. See step 4.1 for additional guidance.

Table E1. Graduation Outcomes Tracker

Binary Predictor	Graduated On Time Number	Graduated On Time Percentage	Did Not Graduate Number	Did Not Graduate Percentage	Total
Indicator: Attendance					
		%		%	
		%		%	
Total		%		%	
Indicator: Behavior					
		%		%	
		%		%	
Total		%		%	
Indicator: Course Performance					
		%		%	
		%		%	
Total		%		%	
Indicator: Composite/Other					
		%		%	
		%		%	
Total		%		%	

Note. An editable version of this table is available in *District Guide for Creating Indicators for Early Warning Systems — Editable Planning Forms*.

Appendix F. Template for Calculating Performance Measures of Indicators and Identifying Optimal Cut Points Through Receiver Operating Characteristic Analysis (from Section 5)

This appendix includes two tools for districts to compare indicators. Table F1 is the template for calculating performance measures of indicators created based on the different cut points. The cut points could be actual values from district variables or selected values that are at certain percentiles. An introduction to Receiver Operating Characteristic (ROC) analysis that districts with statistical capacity can apply in identifying optimal cut points follows table F1 (on next page).

Table F1. Template for Calculating Performance Measures of Indicators Created Based on Different Cut Points

Potential Cut Points for Number of Course Failures	True Positive	False Negative	True Negative	False Positive	Sensitivity	Specificity	Youden Index
Cut point 1	**TP**	**FN**	**TN**	**FP**	**TP/(TP + FN)**	**TN/(TN + FP)**	**Sensitivity + Specificity—1**
Cut point 2							

Note. An editable version of this table is available in *District Guide for Creating Indicators for Early Warning Systems — Editable Planning Forms*.

Identifying Optimal Cut Points Using ROC Curve Analysis

Districts with statistical capacity can identify optimal cut points for continuous variables through ROC curve analysis. ROC curve analysis can be performed by conducting logistic regression using statistical software packages such as SPSS, SAS, and Stata. Logistic regression reports the likelihood of achieving a designated outcome in a log-odds format.

The ROC curve analysis determines a cut point for a continuous variable that optimizes the classification quality of that variable on a certain outcome of interest (Eng, 2005). The optimal cut points, once identified, can be used to create binary indicators that would classify students into graduation "likely" and "not likely" categories. Cross-tabulations between the predicted likelihood and a particular outcome (for example, dropout) would generate four categories: those who were classified as at risk of dropping out and actually dropped out (true positive), those who were classified as not at risk and actually did not drop out (true negative), those who were classified as at risk but did not drop out (false positive), and those who were classified as not at risk but dropped out (false negative). To determine the optimal cut points, you need to calculate true positive and true negative rates. The true positive rate is the proportion of students who dropped out of high school who were successfully identified as at risk of dropping out. The true negative rate is the proportion of graduates correctly identified as not at risk. Ideally, you would want to identify the optimal cut point that has a true positive and true negative as close to 1 as possible.

The ROC curve analysis is driven by a set of true positive and true negative rates for all the possible cut points that exist in the data. The choice of cut points affects cell counts in the true positive, true negative, false positive, and false negative groups, which in turn affects true positive and true negative rates. All pairs of true positive and (1 – true negative) are plotted in a space defined by an X axis and a Y axis. The upper-left corner is the perfect point that signifies student likelihood of dropping out. The goal is to find a cut point that maximizes the power of the predictor variable in predicting the dichotomous outcome of dropping out (or other graduation outcomes). When all possible pairs of the true positive rate and (1 – true negative rate) are plotted on an X-Y graph, a curved line emerges. In the ROC curve analysis, quantitative summary measures of the ROC curve such as the area under the curve (AUC) are often reported. District staff can compare the AUC between indicators. The larger the AUC is, the more effective an indicator is.

Appendix G. Indicator Criterion Decision Matrix (from Section 6)

Unlike the comparisons done to select cut points, the comparisons between indicators should be based on criteria aligned with local priorities. Use this decision matrix (figure G1) to compare your district's priorities against an indicator criterion. See step 6.1 for examples.

Figure G1. Indicator Criterion Decision Matrix

Priority 1: Target only students who are at the highest risk of dropout	**Indicator criterion:** • Highest sensitivity OR • Identify smallest proportion of students as at risk
Priority 2: Identify and provide services to all grade 9 students at risk of dropout	**Indicator criterion:** • Highly specific OR • Identify a large number of students as at risk
Priority 3: Balance both targeting students at highest risk of dropout AND identifying and providing services to all grade 9 students at risk of dropout	**Indicator criterion:** • Balance measure (for example, Youden index)

Consider ranking the priorities for your district before selecting the indicator criterion that most closely matches that priority. Supposing that priority 1 is most aligned with the district goals, then sensitivity should be used to compare and rank the indicators. Table G1 can be used in the comparison and ranking process.

Table G1. Template for Calculating Performance Measures of Indicators and Indicator Ranking

	True Positive	False Negative	True Negative	False Positive	Sensitivity	Rank of indicators (sensitivity)	Specificity	Rank of indicators (Specificity)	Youden Index	Rank of indicators (Youden Index)
Name of indicators	TP	FN	TN	FP	TP/(TP + FN)	Rank	TN/(TN + FP)	Rank	Sensitivity + Specificity—1	Rank
Indicator 1 (for example attendance less than 90%)										
Indicator 2 (for example, failed a core course)										
Indicator 3 (for example, GPA below 2.0)										
Indicator 4 (for exam-ple, being suspended)										
Indicator 5 (for example, failed math)										
....										

Note. An editable version of this table is available in *District Guide for Creating Indicators for Early Warning Systems — Editable Planning Forms.*

REL West website
http://relwest.WestEd.org

email
relwest@WestEd.org

@REL_West

866.853.1831

730 Harrison Street
San Francisco, CA 94107-1242

REL WEST | WestEd