6-1-2010

The Predictive Validity of the Early Warning System Tool

Evelyn Johnson
Boise State University

Carrie Semmelroth
Boise State University

DOI: 10.1177/0192636510380789

The Predictive Validity of the Early Warning System Tool

Evelyn Johnson and Carrie Semmelroth
Boise State University

Abstract
The Early Warning System (EWS) is a tool developed by the National High School Center to collect data on indicators including attendance, GPA, course failures and credits earned. These indicators have been found to be highly predictive of a student's likelihood of dropping out of high school in large, urban areas. The EWS tool was studied in two suburban schools. With the exception of attendance data, findings suggest that the indicators and suggested threshold for risk determination are predictive in suburban contexts.

Keywords: drop-out, at-risk, screening, RTI

This research examines methods of screening at the secondary level (grades 6-12) within an RTI framework. The RTI framework incorporates a tiered instructional service delivery model, rooted in the preventive sciences, that has as its ultimate goal school improvement across the K-12 spectrum (Mellard & Johnson, 2008). In a preventive model of service delivery, schools take a proactive approach towards the early identification of students at-risk for poor academic outcomes. Therefore, a critical component within the RTI framework is the development of a screening process that helps determine which students are at-risk for not achieving the primary outcomes of interest (Johnson, Smith & Harris, 2009). A critical component within the screening process is identifying and defining the primary outcomes of interest.

At the elementary school level, the primary academic outcome of interest is the development of academic skills that will support success in later school years. But the goals and outcomes at secondary levels are quite different from those at elementary levels, and may not be consistent across all secondary students (Johnson et al, 2009). For example, some students will attend a four-year college or university. Others will seek vocational or technical training. Others may attend a local community college to continue preparation for a four-year institution of higher learning. Still others will immediately join the workforce or armed services. The variety of outcomes for which we are preparing students makes it difficult to know what we would be trying to predict with a screening process.

Though long-term goals may vary, for all students, obtaining a high school diploma is a shared short-term outcome. High school dropouts face significantly higher probabilities of incarceration, poverty and a need for social services (Schweinhart, 2004). While districts and states differ on the specifics of high school graduation requirements, most include a combination of successful course and credit completion, successful performance on exit exams, and other requirements such as senior projects. These requirements help provide a common system of evaluation for all students and suggest that screening efforts should be directed to identify the following groups of students:

1. Students who are at risk for dropping out of school.
2. Students who have learning needs that require sustained intervention.
3. Students who are at risk for not meeting performance benchmarks on grade-level state assessments (Johnson et al., 2009).

The focus of the current study is on the identification of the first group of students, those at risk for dropping out of school. In order to devise an effective RTI system that can support students, a universal screening system at the secondary level needs to identify early (as early as 9th grade) those students most at-risk for dropping out of school. Effective screening tools are characterized as brief and efficient measures that accurately sort students into one of two groups: (a) those who are at risk for the outcome of interest and (b) those not at risk (Jenkins, 2003). Screening measures must include the specific indicators that are the best predictors of determining students at-risk for eventually dropping out of school. Research on high school drop-out rates indicates four patterns that emerge as early as sixth grade that predict a higher risk of students dropping out of high school:

1. Academic performance track – Students who have a pattern of low grades, who have low test scores, who are failing core courses and who fall behind in course credits.
2. Engagement track – Students who have high rates of absenteeism, poor behavior records and bad relationships with teachers and peers.
3. Combined academic and engagement track – Students who experience difficulties both tracks.
4. Transition years track – Students who experience sharp declines in performance as they transition from elementary to middle school and then from middle to high school (Jerald, 2006).

Large scale studies conducted in urban areas such as Philadelphia and Chicago have confirmed these tracks (Jerald, 2006). These studies have also indicated the importance of the occurrence of "high yield" indicators during the high school transition year such as attendance, grade point average, and the number of courses failed as predictors of dropping out of school (Heppen, O'Cummings, & Therriault, 2008; Jerald, 2006; Neild, Balfanz & Herzog, 2007). The transition year from middle to high school (typically ninth grade) has been found to be the "make or break" year for predicting high school completion (Heppen, et al., 2008). The transition year experience contributes substantially to the probability of dropping out, despite controls for demographics and previous school performance (Neild et al., 2007). Two of the most powerful predictors of high school completion identified in the literature to date are attendance and course performance from the first year of high school (Allensworth & Easton, 2007).

Based on the research identifying these high yield indicators, the National High School Center developed an Early Warning System (EWS) tool for the first-year transition to high school using student data based on attendance, course performance (based on GPA, courses completed and failed) and the "on-track" indicator (a combination of course failures in core academic course and credits earned) (Heppen, et al., 2008). The EWS was developed for schools and districts to automatically calculate indicators related to attendance and course performance in order to identify if a student is at risk for dropout, or on-track for graduation.

The EWS uses the first-year of high school data to identify students at risk using "highyield" indicators (see Table 1). The "high-yield" indicators are research-based, strong, early warning signs that predict if a student will graduate from high school. Table 1 summarizes the indicators that are highly predictive of student outcomes (Heppen, et al., 2008). There are four different benchmarks that "red flag" a student that may be at risk for dropout including: (a) missing more than 10% of instructional time during the first year, (b) missing more than 10% of the first 20 days, (c) earning a GPA under 2.0, and (d) failing one ore more courses. Emerging evidence confirms the predictive power of these indicators. Low attendance during the first 20 days of ninth grade were found to be a more powerful predictor than other existing data such as test scores, age or academic failure (Jerald, 2006). Given that attendance is routinely collected in schools, information about absences may be the most practical indicator for identifying students in need of early interventions (Allenworth & Easton, 2007; Heppen, et al., 2008; Jerald, 2006).

The "on-track" indicator requires that students have no more than one failing grade per semester, and no fewer than the number of credits required to be advanced to the 10th grade. "On-track" essentially reflects the bare minimum performance for a student at the end of her/his first year in high school. The minimum performance equals one-fourth the total number of credits required for graduation, minus one. Students who are identified as "off-track" at the end of their first year in high school should be considered at risk for dropout and should be targeted for intervention (Heppen, et al., 2008). The on-track indicator is emerging as a quality measure to predict high school outcomes. For example, the on-track indicator was 85% successful in predicting which members of a freshman class in Chicago would not graduate (Jerald, 2006).

It is important to note that these on-track and high yield indicators have been studied in heavily urbanized contexts such as Philadelphia and Chicago (Heppen, et al., 2008; Jerald, 2006; Neild et al., 2007). As Jenkins (2003) notes, differences in the local context may dictate different predictors, different cut scores or different approaches to screening. Therefore, screening tools and suggested guidelines require cross-validation across settings. At the time of this writing, there are no published cross-validations establishing the effectiveness of the EWS tool in suburban or rural areas. The purpose of the current study was to determine the predictive validity of the EWS tool high yield indicators within settings other than large urban areas. The current study is part of a larger research project designed to develop a screening procedure for use within a secondary level RTI framework.

DOI: 10.1177/0192636510380789

Methods

Setting and Participants

Two high schools in the Northwest were selected to participate in this study. The first high school (HS1) contains grades 10 through 12 and is the only high school within a small suburban district. Only 20% of the district population is eligible for free or reduced lunch (FRL). The second high school (HS2) contains grades 9 – 12 and is located in a suburban area, in the third largest school district in the state. Over 50% of the district population is eligible for FRL. Demographics of both schools and districts are provided in Table 2. Because the purpose of this study was to test the use of the EWS tool within a high school, data was not aggregated across schools. Therefore, information on participants, procedures and results are presented by school.

Participants in High School 1. For HS1, student records for the graduating class of 2008 were collected from the school registrar. For the 2007-08 school year, there were 243 total seniors. Because this is a grade 10-12 high school, the tenth grade year was selected as the transition year. Working backwards from their sophomore year (2005-06 school year), databases containing relevant information for the EWS tool were constructed. Of the initial 243, 40 lacked sophomore data because they had transferred in from different schools, and 3 lacked confirmed graduation information, leaving a sample of N = 200 for HS1. The removal of the 43 students from the analysis did not alter the proportional demographic make-up of the sample.

Participants in High School 2. For HS2, student records for the graduating class of 2008 were collected. For the 2007-08 school year, there were 246 total seniors. This high school contains grades 9-12, so databases were constructed from their freshman year data (2004-05 school year). Of the initial 246, 59 lacked freshman data, and 15 lacked confirmed graduation information leaving a sample of N = 172 for HS2. The removal of the 74 students from the analysis did not alter the proportional demographic make-up of the sample.

Protecting participants' rights. The data was collected through the use of extant data bases using unique student identifiers without names attached. Follow-up inquiries were directed to school personnel through the use of these student identifiers.

Procedures and Data Collection

The EWS tool developed by the National High School Center was used to collect data for this study. The EWS tool is a formatted Microsoft Excel workbook that contains several linked worksheets for analyzing student data and is available for free download from the National High School Center's website, www.betterhighschools.org. Based on the data entered into the EWS tool, a flag of "Yes" or "No" appears under the specific indicator. For example, if a student has a GPA of less than 2.0, a "Yes" appears in that cell on the summary sheet of the EWS tool. After all data is entered, an omnibus indicator in the form of a red flag appears on the spreadsheet. A student who has a "Yes" for any one or more of the indicators also has a red flag appear in the left most column of the Excel database immediately next to their name, providing a quick, visual indicator for the user.

The predictor variables included the high yield indicators (absentee rate, GPA, course failures, on-track indicator) identified by Heppen and Therriault (2007). Data collection began with a review of the students' records from their transition year of high school. For HS1, we used the sophomore year data and for HS2 we used freshman year data, as transition years have been shown to be critical time periods for screening for at risk behaviors (Jerald, 2006). Data on these variables were collected for both schools from a variety of sources. Attendance records were assembled from the school registrar records. Absences were counted manually and recorded in half day and full day increments. This information was entered into an Excel spreadsheet. Another school database, Power School, was used to collect information on credits earned and grade point average (GPA). Grade Point Average (GPA) and credits failed (core and total) were exported from PowerSchool and input manually into the Excel spreadsheet. From this data, the information for the On-Track indicator was computed. In cases of incomplete records, further information was gathered through an interview with the assistant principal, registrar and school psychologist. All data on the predictor variables were then imported to the EWS spreadsheet and a master list of students was created.

Outcome Variable

Successful completion of high school was the outcome variable and for the purposes of this study was defined as the student successfully completing high school either in the 2008 or 2009 school year. Graduation data was obtained from the registrar in both schools. The State's Department of Education most current School Report Card cites the current high school completion rate for 2007 at 88% (State Department of Education, 2008). The official definition

DOI: 10.1177/0192636510380789

used by the state is: “A drop out is a student who was enrolled some time during the current year but was not enrolled at the end of the current regular school year; or was enrolled at the end of the prior regular school year and expected to be part of the membership of the current school year but did not enroll in the current school year; and has not graduated from high school or completed a state or district approved educational program and does not meet any of the following exclusionary conditions: (a) transfer to another public school district, private school, or other state or district approved program; (b) temporary school-recognized absentee due to suspension, illness or death” (State Department of Education, 2008).

Data Analysis

The analysis of data examined the efficacy of each of the single predictors as well as the overall “red flag” that the EWS tool assigns to a student record that contains a flag for any reason. Screening data can be evaluated using a number of statistics, each of which provides different information about the measure’s potential utility as an accurate classification tool. For each of the predictor variables, we computed the classification accuracy, sensitivity, specificity, Receive Operating Characteristic (ROC) curve and relative risk estimate. Each of these statistics is explained below.

Classification accuracy is a measure of how well a screen accurately sorts students into at-risk or not at-risk categories (Jenkins, Hudson, & Johnson, 2007). It is calculated as the number of students who are accurately classified divided by the total number of students. Sensitivity is the probability that a screening test will be positive when the student is at-risk. Specificity is the probability that a screening test will be negative when the student is in fact not at-risk. The sensitivity and specificity of screening measures should be high if a measure is to be used as part of an efficient and effective screening process within an RTI framework, with some researchers advocating for sensitivity and specificity levels as high as 90% (Compton, Fuchs, Fuchs & Bryant, 2006; Jenkins, 2003).

Many of the statistics associated with screening data are calculated under a specific condition or classification rule. Different rules result in different levels of sensitivity, specificity and classification accuracy. Therefore, it is often helpful to have a way of displaying and summarizing performance of a screening measure over a wide range of conditions (Krzanowski & Hand, 2009). Receiver-operating characteristic (ROC) plots graphically display the set of potential combinations of sensitivity and specificity possible for predictors (Pepe, Janes,Longton, Leisenring, & Newcomb, 2004) and provide a useful way of determining a screen’s potential as a classification tool (Zweig & Campbell, 1993). Instead of providing data only on one predetermined set of cut scores, the ROC curve provides a graphical display of a measure’s classification accuracy across all cut scores. An overall indication of the diagnostic accuracy of a ROC curve is the area under the curve (AUC). The AUC values closer to 1 indicate the screening measure reliably distinguishes among students with satisfactory and unsatisfactory reading performance, whereas values at .50 indicate the predictor is no better than chance (Zhou, Obuchowski, & Obuchowski, 2002).

Finally, because many screening statistics are criticized for not being helpful in clinical applications, we also computed the relative risk estimate of each indicator. The relative risk estimate compares the probability of an outcome in each group (Garson, 2009). For each group a risk ratio is computed, and the relative risk is the ratio of the two ratios (Garson, 2009). The relative risk estimate has been described in medical research as more consistent with the way people think about the probability of an event occurring given a specific condition (Garson, 2009), and in the case of the current study, gives the practitioner an indication of the meaning of having a specific flag. In this study, we calculated relative risk using the following procedure. First, the risk of dropping out for a student with the indicator present was calculated. Then the risk of dropping out for a student without the indicator present was calculated. The relative risk is then calculated as the ratio of these two ratios. The interpretation of the relative risk estimate is explained in more detail in the discussion section.

In order to conduct these analyses, data files from the EWS tools were exported into SPSS and merged with the graduation rate data. All data analysis was conducted using SPSS version 17.

Results

Table 3 presents the classification accuracy, sensitivity, specificity, ROC AUC and relative risk estimate for each of the predictors by school. In this section we present the results by school. Findings across schools are presented in the discussion section of this manuscript.

High School 1

There were a total of 23 dropouts at HS1, which represents 12% of the sample, and is consistent with the state's reported drop out rate. As presented in Table 3, attendance had the highest classification accuracy of the single predictors, followed by the off-track indicator and GPA. However, attendance also had one of the lowest sensitivity levels, identifying only 36% of the students who dropped out. Of the single predictors, GPA had the highest sensitivity level (84%), with a high associated level of specificity (86%). The number of course F's received had the next highest sensitivity level (78%) with an associated specificity of 84%. Given that GPA and course Fs are related to one another, it is not surprising that they would result in similar levels of sensitivity and specificity. The number of course Fs had the highest ROC AUC (.87)
followed by GPA (.85).

As indicated in Table 3, the presence of any one of the risk indicators significantly reduced a student's chance of successfully completing high school. Students with a red flag for any reason had a 2% probability of successfully completing high school. Even the least effective predictor (first 20 days absences) however, resulted in a relative risk estimate of 19%, with other indicators ranging from 5 to 9%. The relative risk estimate and its interpretation are discussed in more detail in the discussion section.

The 'red flag' indicator resulted in the lowest classification accuracy (77%), primarily because of the high false positive rate. However, the red flag resulted in the highest sensitivity levels (96%), identifying 22 of 23 students who dropped out. The AUC for the red flag indicator was also high (.85), ROC AUCs in this range are considered to have good predictive utility (Krzanowski & Hand, 2009).

High School 2

There were a total of 15 confirmed students who dropped out in HS2, representing 9% of the sample, which is slightly below the reported state rate of 12%. Of the single predictors, the off-track indicator resulted in the highest classification accuracy (88%), followed by GPA (82%) and attendance (81%). The number of course F's had the highest sensitivity levels (87%) with an associated specificity of 69%. GPA also had a high sensitivity level (86%) but a higher specificity of 81%. Attendance resulted in the lowest sensitivity (27%), identifying only 4 of the 15 students who dropped out. GPA was the only single predictor that resulted in a ROC AUC greater than .80. The red flag indicator resulted in the lowest classification accuracy of all indicators (67%), primarily due to the overidentification of 56 students. However, the red flag also identified all of the students who eventually dropped out, resulting in 100% sensitivity.

The relative risk estimates of the single predictors varied from 5 to 47%. For this high school, attendance emerged as a risk factor, but was not as strong as those indicators related to course performance. No risk estimate for the red flag could be computed because one of the cells contained zero students.

Discussion

The purpose of this study was to determine the accuracy of the high yield indicators that comprise the EWS tool in suburban contexts and smaller high schools. Most of the research on these indicators has been conducted in large, urban areas such as Philadelphia and Chicago. In this study, we tested the power of these predictors in very different settings and across different populations. Additionally, the studies to date examining the utility of the predictors have been conducted with freshman year data, although the literature discusses the importance of a transition year. In the present sample, HS 1 includes only grades 10 – 12, making tenth grade the transition year for this school, as well as all high schools that include grades 10 -12. In order to determine whether the EWS tool should be used to make important school and student level decisions about intervention, it is important to collect evidence about the validity of the predictors within the context of its use (Jenkins, 2003). The current study provides additional evidence of the predictive power of the high yield indicators in settings other than large urban areas. In this section we a) discuss our findings and their consistency with studies conducted to date, b) discuss limitations of the current study to include general concerns about the methods used to measure some of the variables in the EWS tool, and c) discuss the implications for practice.

Consistency of Findings with Other Studies

For both high schools, the red flag (omnibus indicator) captured all or nearly all (sensitivity of 96% and 100% in HS1 and HS2, respectively) of the students who dropped out. In HS1, specificity of the red flag was high at 75%, but in HS2, specificity was only 64%. In student counts, these specificity levels correspond to 44 false positives in

HS1 and 56 false positives in HS2. Guidelines for acceptable levels of sensitivity and specificity vary depending on the outcome we are trying to predict. In many cases, practitioners would counsel to err on the side of caution – over identification is preferable to letting a student who is truly at risk fall through the cracks. Of course, this has to be tempered by resource availability. Most schools operate with limited resources, and many research-based interventions designed to support students who are at-risk for poor academic outcomes have been found effective when implemented in small teacher:student ratios. Given the importance of the outcome in question (dropping out from high school), the rates of overidentification found in this study may constitute an acceptable level of error for practitioners. Our findings on the overall efficacy of the combined use of the high yield indicators are consistent with summaries of research conducted across various settings (see for example, Allensworth & Easton, 2005; Allensworth & Easton, 2007, Jerald, 2006; Neild & Balfanz, 2006; Rumberger, 2004; Roderick, 1993).

Of the single high yield predictors, GPA tended to be the strongest predictor in both high schools. Attendance was not as strong of a predictor in either school as was found in prior studies in Philadelphia and Chicago (Jerald, 2006). We conjecture that one explanation for this difference in findings comes from the challenges we encountered in the participating schools with reconstructing attendance records. In both schools, attendance is based on teacher data input to the school data system. The reliability of the attendance data in this study is unclear as we detected inconsistencies during the transfer of attendance data from school records to our database. Therefore, although our findings suggest that attendance is not a strong predictor of drop-out rates, we caution that our findings for attendance may not be reliable, and that the difficulties we encountered indicate the need for schools to ensure the accuracy of the way in which they collect data on their predictor variables.

Limitations

Screening research is especially susceptible to error due to a variety of factors primarily related to sampling procedures and measurement error. Like many screening studies therefore, the results of this study must be interpreted with caution. First, our results are based on one cohort’s worth of data for each high school. Cross-validation is a critical part of screening research – if results are not consistent across multiple cohorts, then it is difficult to have confidence in the decision rules developed based on the results of the analysis. Cross validation studies also give confidence in our screening instruments if findings are consistent across samples and settings. In this regard, although we only have one cohort from each school, the current study does represent a cross-validation of the high yield indicators that have been identified within urban contexts. With the exception of our findings on attendance, our results are consistent with these studies. Ideally however, as recommended by Jerald (2006), decision rules about the high yield indicators should be verified across multiple cohorts within a school so that the school can refine their decision rules.

An additional concern with this study is the way in which attendance and drop out data were recorded. The accuracy of the attendance data in both high schools is in question because of the way in which it is monitored at both schools. Teachers take attendance and enter the data into the school data system, Power School. The by class period attendance data is then aggregated and a number of days absent is calculated. In both schools, as we reconstructed records for students who were not in the system, we noted error patterns in attendance record keeping. For example, a student who was marked as absent for multiple days by most teachers was also marked as consistently attending one class in the middle of the day. A cross-check of this teacher’s records verified that attendance data was input only intermittently. Similar issues occurred in both high schools. It is unclear the extent to which a lack of validity of our attendance data impacted its ability to serve as a useful predictor of drop-out rate. In both schools, attendance failed to identify students who later dropped out (too many false negatives). This indicates that the current threshold for flagging students for poor attendance (missing > 10% of days within a semester) may be set too high for the two schools in the current study. Changes to the threshold however, cannot be considered until attendance is collected with greater precision.

The validity of our drop out rate is also in question. For both schools, there were a number of students for whom we were missing high school outcome data. Because this study was conducted in a state that does not issue unique student identifiers state-wide, when a student transfers districts, they receive a new student identification number. Similarly, if a student moves out of state, the data for that student is not linked – we have no efficient way to determine if that student successfully completed school. For the 99 students for whom we have no outcome data, we were unable to determine whether they had transferred or dropped out. The difficulty of getting an accurate drop-out rate is a nationally recognized problem (Jerald, 2006). In the context of this study, the difficulty of obtaining an accurate dropout rate may have impacted some of our analyses.

DOI: 10.1177/0192636510380789

Implications for Practice

One of the difficulties with screening research is that the results of the statistical analyses are not easily translated to practice. For example, what does a sensitivity level of 50% suggest a practitioner should do when confronted with a student who presents as at-risk on that indicator? What is an acceptable level of overidentification compared to missing students who are at-risk? These types of cost-benefit determinations can be difficult to calculate, which is part of the challenge of developing decision rules for screening. For example, if we use classification accuracy as the guideline for decision making, in both schools, the Off-Track indicator (course credits and course F's) was a more accurate predictor than the Red Flag. However, the sensitivity level of the off-track indicator was low (50% in HS1, and 67% in HS2), missing significant numbers of students who dropped out. Attendance performs similarly with high classification accuracy (92% and 81% respectively) but low sensitivity levels (36% and 27% respectively). Of the single predictors, a student's GPA appears to offer high classification accuracy with near optimal levels of sensitivity and specificity.

The ROC curves provide additional evidence of these findings. In both schools, GPA and the red flag indicator had the highest AUC, though no predictor reached an AUC of > .90, which is generally considered to be excellent. Similarly, the ROC AUC statistics for attendance are poor, an AUC of .5 is obtained by chance, and the .66 and .58 obtained for HS1 and HS2 respectively are not much higher than this number.

ROC curves and sensitivity levels are useful ways to discuss the accuracy of screening measures in general terms. Such analyses inform research efforts that subsequently inform policy decisions. However, they are not very useful statistics in practical terms – for decision making at the individual student level. Knowing that earning a GPA of less than 2.0 has a sensitivity level of 86% is probably not very informative for a teacher. For this reason, we included the relative risk estimate as a part of our screening analyses. The utility of the relative risk estimate as a way to interpret screening data by practitioners is demonstrated by this study. For example, in High School 1, a student with a flag for missing the first 20 days of school has a relative chance of 19% of completing high school compared to students without that flag. Students who have a red flag have a 2% relative chance of completing high school successfully compared to students without a red flag. In other words, this is a student whose performance indicators warrant closer scrutiny. For all the indicators except for those related to attendance, a student with a flag in any area has a probability of less than 10% of graduating from school, when compared to a student without those indicators.

If important decisions are to be made about student performance and interventions, then it also seems that another important implication for practice is that schools need to provide training and support to prevent the Garbage In, Garbage Out (GIGO) principle. Inaccurate estimates of our predictor variables clearly limit our ability to use them to make decisions. Similarly, any tool that attempts to predict whether a student will drop out of school needs to be based on accurate drop out information. The simplicity of the EWS tool lies in its reliance on data that is already collected on students and that is highly predictive of successful completion of high school. Its implementation could help drive the development of better student tracking systems first at the state level (already in progress in many states), and ultimately, at the national level. When more accurate statistics about the prevalence of a condition can be obtained, we can obtain more accurate statistics about the variables that predict that condition. Finally, the results of our study demonstrate highlight the need for school districts to become adept at the type of analysis presented in this article, because ultimately, rules about cut scores and benchmarks will likely be developed within the local context.

While the EWS tool alone does not constitute a full screening for potential academic and behavior concerns, it does provide schools with an essentially no cost system for screening and provides important school level feedback. For example, if a school finds that a significant number of students have a flag in the "off track" indicator at the end of their freshman year, they might need to in the short term, consider a system of support for credit completion but also have a system of screening measures in the academic areas to help identify in which courses and areas a system of interventions might be required. For example, within the larger study of which this study is a part, we are examining the predictive validity of a number of screening measures and attempting to integrate that data within the EWS tool to facilitate data storage, analysis and decision making, particularly in schools with limited resources.

DOI: 10.1177/0192636510380789

Finally, the RTI framework is predicated on an early identification and intervention model. Although the research on the high yield indicators suggests that transition years (e.g. the first year of high school) are the critical times to evaluate these factors, in high schools that include only grades 10-12, this may be problematic. Further research on the EWS tool should investigate its predictive validity as early as middle school. Can we accurately identify students in 6th, 7th or 8th grade who are at increased risk for dropping out? Early identification allows for the development and implementation of interventions that may be more successful than waiting until high school to support students. The movement of the EWS to middle school might also facilitate a stronger dialogue between middle and high schools, and ultimate allow for district level policies that provide a complete spectrum of support for students.

Conclusion

This study examined the predictive validity of the high yield indicators that comprise the EWS tool in predicting whether a student is at-risk for dropping out of school. Given the consistency of the findings of this study with those conducted in large, urban areas, our findings provide support for the use of these indicators in suburban settings, and highlight the need for developing more consistent and accurate methods of measuring the drop out rate as well as other indicators such as attendance. The construction of a database such as the EWS to collect this information is an initial step for high schools interested in implementing a preventive model of service delivery that proactively identifies, and intervenes early for students at-risk of dropping out of school.

References

Allensworth, E. & Easton, J. Q. (2005). *The on-track indicator as a predictor of high school graduation.* Chicago: Consortium on Chicago School Research.

Allensworth, E. & 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: Consortium on Chicago School Research.

Compton, D. L., Fuchs, D., & Fuchs, L. S., & Bryant, J. D., (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology, 98,* 394-409.

Garson, G. D. (2009). Log-linear, logit, and probit Models, from *Statnotes: Topics in Multivariate Analysis.* Retrieved 10/15/2009 from http://faculty.chass.ncsu.edu/garson/pa765/statnote.htm.

Heppen, J., O'Cummings, M., & Therriault, S. B. (2008). *Early warning systems tool*. Washington, DC: National High School Center at the American Institutes for Research. Retrieved from http://www.betterhighschools.org/pubs/#Dropout

Heppen, J. B. & Therriault, S. B. (2008). *Developing early warning systems to identify potential high school dropouts*. Washington DC: National High School Center.

Jenkins, J. R. (2003, December). Candidate measures for screening at-risk students. Paper presented at the NRCLD responsiveness-to-intervention symposium, Kansas City, MO. Retrieved April 3, 2008, from http://www.nrcld.org/symposium2003/jenkins/index.html

Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for service delivery in an RTI framework: Candidate measures. *School Psychology Review, 36*, 582-99. Jerald, C. D. (2006). *Identifying potential dropouts: Key lessons for building an early warning data system*. Washington, DC: Achieve, Inc. American Diploma Project Network.

Johnson, E. S., Smith, L., & Harris, M. L. (2009). *How RTI works in secondary schools.* Thousand Oaks, CA.

Krzanowski, W. J. & Hand, D. J. (2009). *ROC Curves for Continuous Data.* Boca Raton, FL: CRC Press.

DOI: 10.1177/0192636510380789

Mellard, D. F. & Johnson, E. S. (2008). *RTI: A practitioner's guide to implementation.* Thousand Oaks, CA: Corwin Press

Neild, R. C. & Balfanz, R. (2006). An extreme degree of difficulty: The educational demographics of urban neighborhood high schools. *Journal of Education for Students Placed at Risk, 11*(2), 123-141.

Neild, R.C., Balfanz, R., & Herzog, L. (2007). An early warning system. *Educational Leadership, 65*(2), 28-33.

Pepe, M., Janes, H., Longton, G., Leisenring, W. & Newcomb, P. (2004). Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *American Journal of Epidemiology, 159,* 882-890.

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

Rumberger, R. W. (2004). Why students drop out of school. In Orfield, G. (Ed.), *Dropouts in America: Confronting the graduation rate crisis* (pp. 131-156). Cambridge, MA: Harvard Education Press.

Schweinhart, L. J. (2004). *The High Scope Perry preschool study through age 40: Summary, conclusions, and frequently asked questions.* Ypsilanti, MI: High Scope Press.

Zhou, X. H., Obuchowski, N. A., & Obuchowski, D. M. (2002). *Statistical methods in diagnostic medicine.* Wiley & Sons: New York.

Zweig, M. H. & Campbell, G. (1993). Receiver-Operating Characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clinical Chemistry, 39* (4), 461-577.

DOI: 10.1177/0192636510380789

Table 1

Summary of "High-Yield" Indicators

Type of Information	Indicator	Brief Description	"Red flag" (indicates a student may be at risk for drop out)
Attendance	Absenteeism Rate	Number of days absent during the first 20 days and each quarter of the first year in high school	More than 10% of instructional time missed during the first year
Course Performance	Course failures	Number of Fs in any semester-long course during the first year in high school	One failed course
	GPA	GPA for each semester and cumulative GPA	GPA under 2.0
	On-track indicator	Combination of number of Fs in core academic courses and credits earned during the first year of high school	Two or more Fs in core academic courses and/or fewer than one-fourth of the credits required graduate minus one

Table 2
District and school demographics.

HS 1		HS 2		
	District	School	District	School 1267
Total Enrollment	2683	578	14681	13%
English Language Learners	<1%	<1%	17%	53%
Free/Reduced Lunch	20%	16%	55%	12%
Special Education	11%	5%	14%	66%
White	90%	91%	64%	31%
Hispanic	3%	3%	29%	1%
Asian	3%	2%	3%	<1%
Black	2%	2%	2%	<1%
Other	2%	2%	2%	

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at The Predictive Validity of the Early Warning System Tool, published by SAGE. Copyright restrictions may apply.
DOI: 10.1177/0192636510380789

Table 3
Classification accuracy, sensitivity, specificity, ROC AUC and odds ratios of predictors by school.
Note. CA = Classification Accuracy. Sen = Sensitivity. Spec = Specificity. ROC AUC =Receiver

Predictor	CA	Sen	Spec	ROC AUC	Relative Risk	TP	TN	FP	FN
				High School 1					
First 20 days	88	35	94	.64	.19	8	167	10	15
GPA	89	84	86	.85	.05	16	151	25	3
Course F	84	78	84	.87	.08	18	149	28	5
Off track	91	50	96	.73	.09	10	170	7	10
Attendance	92	36	99	.66	.08	8	176	1	15
Red Flag	77	96	75	.85	.02	22	133	44	1
				High School 2					
First 20 days	80	47	83	.63	.28	7	130	27	8
GPA	82	86	81	.83	.05	12	125	29	2
Course F	70	87	69	.78	.09	13	108	49	2
Off Track	88	67	90	.78	.08	10	142	15	5
Attendance	81	27	87	.58	.47	4	136	21	11
Red Flag	67	100	64	.82	Note*	15	101	56	0

Operating Characteristic Area Under the Curve. TP = True Positives. TN = True Negatives. FP = False Positives. FN = False Negatives. *No value could be computed for the relative risk because one of the cells contained a 0 value.