Is GPA or ACT Composite more related to student on-time graduation in NDUS Community Colleges?

The data set used to examine this question was made up of students who graduated from a North Dakota high school between 2008 and 2011 and were enrolled in an NDUS Community College immediately following graduation. Any student who was identified as having transferred to an NDUS Regional University, an NDUS Research University, or an out-of-state University was not considered in this study. This was done in an attempt to isolate this study as much as possible to students whose objective was to earn an Associate Degree, as opposed to those that attend an NDUS Community College with the intention of transferring credits before the completion of an Associate Degree. For the purposes of this study, "on-time" completion was considered to be a student completing his or her Associate degree within 3 years of initial, full-time enrollment in an NDUS Community College. The total sample size was 2,351 students.

	ACT Composite	HS GPA	On Time Graduation
ACT Composite	1.000	0.540	0.161
HS GPA	0.540	1.000	0.344
On Time Graduation	0.161	0.344	1.000

Table 1: Correlation Matrix for ACT, GPA, and On-Time Graduation

To begin, a correlation matrix was constructed to determine if high school GPA (on a 4.0 scale) or ACT Composite was linearly correlated with on time completion of an Associate degree. As can be seen in Table 1, the high school GPA of the students in this study was more strongly correlated with on-time completion, with a correlation coefficient of 0.344, compared to that of 0.161 for ACT Composite and on-time degree completion.

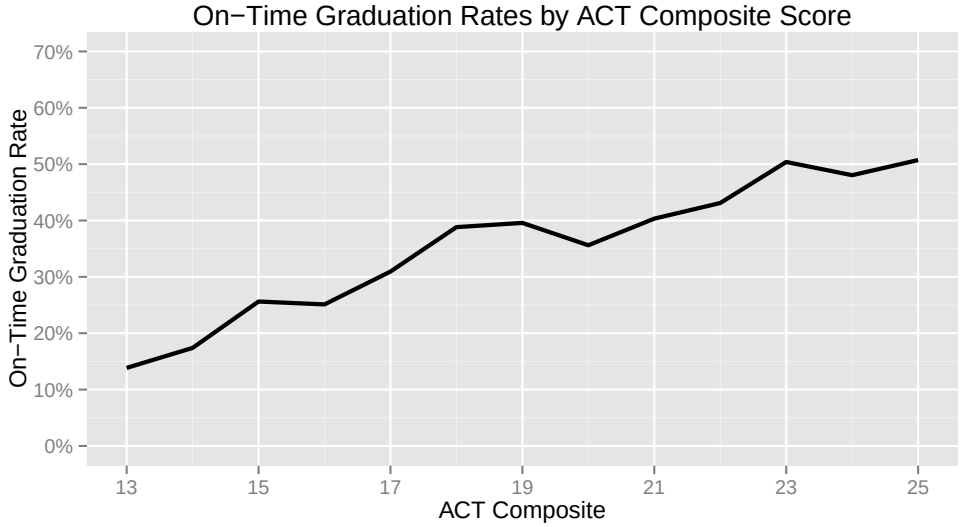

Figure 1: On-Time NDUS Graduation Rates by ACT Composite Score

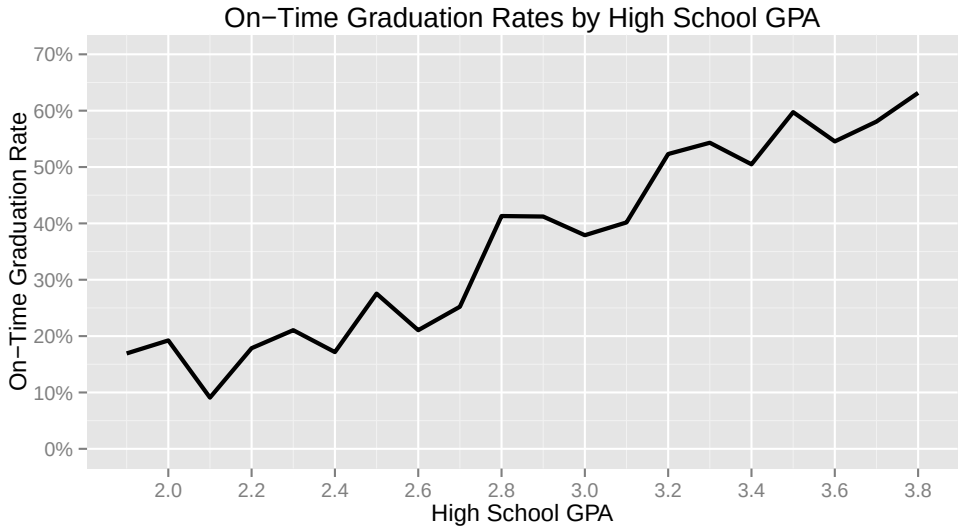

Figure 2: On-Time NDUS Graduation Rates by High School GPA

As Figures 1 and 2 show above, both ACT Composite score and High School GPA are positively correlated with an increase in on-time graduation rates in students. (Please note the figures only show values for cases with 50 or more students)

To further determine their relative usefulness in assessing the likelihood of an incoming freshman to graduate, logistic regression models were developed using an indicator variable for on-time graduation with ACT Composite score or High School GPA as the independent variables. Prior to the development of the models, the student sample was split into two random halves, one to develop the model, and one to test it. After determining the two logistic regression models used to assess student likelihood of on-time Associate degree completion using the development sample, receiver operating characteristic (ROC) curves were created for each model using the test sample to determine the relative value of each.

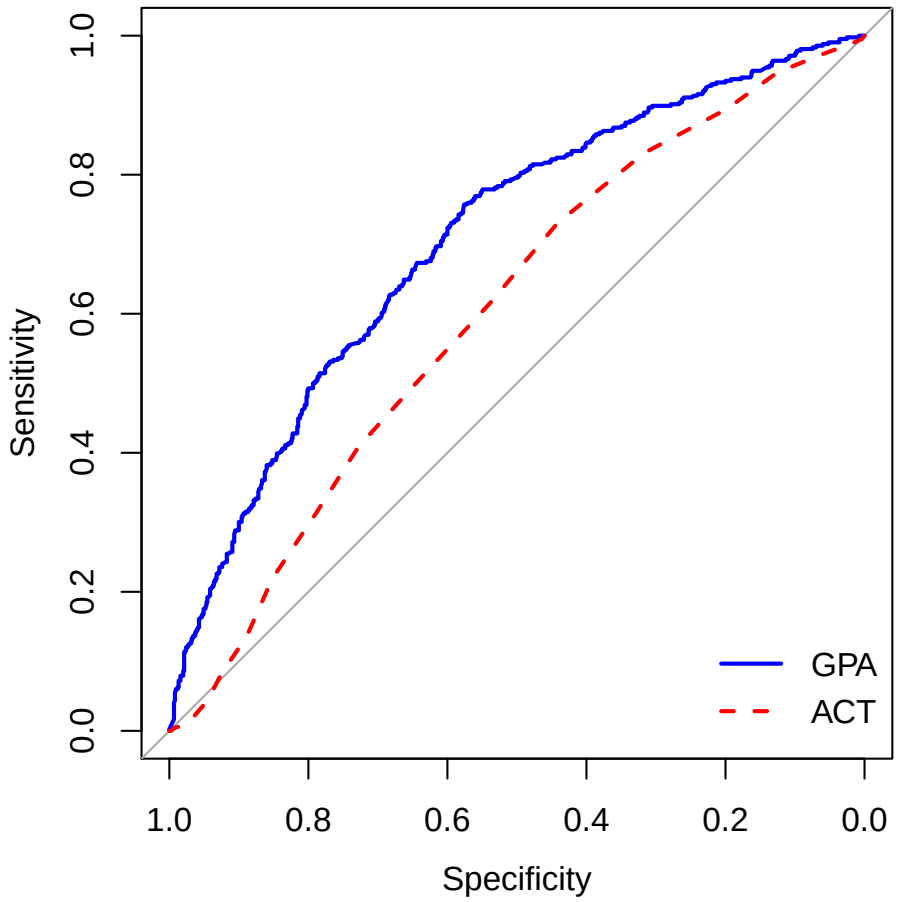

Figure 3: GPA vs. ACT Composite ROC Curves

ROC curves measure the tradeoff between the sensitivity (true positive rate) and specificity (true negative rate) of a given model. Models with a higher level of accuracy will have an ROC curve closer to the upper left hand corner of the graph, indicating the model is accurately identifying both on-time graduates and students who do not graduate on-time at equally high rates. To help determine the relative usefulness of models plotted on an ROC curve, the area under the curve, or AUC is calculated. To interpret the AUC, suppose that the two groups of students were already pre-split, with a group of on-time graduates, and a group of students who did not graduate on-time. The AUC represents the probability that, when presented with a randomly selected pair of students, one from each group, the model will correctly assign a higher probability of success to the on-time graduate. Very poor models will have an AUC of close to 0.50, with excellent models having an AUC close to 1.

As shown in figure 3, the AUC for the logistic regression model utilizing GPA is 0.706. The AUC for the logistic regression model using ACT Composite Score is 0.602. The p-value for the significance test to determine if a difference exists between the two AUC's was $p < 0.0001$, indicating that the logistic regression model utilizing GPA was statistically significantly more useful at identifying the probability of student on-time Associate degree completion than was the model using ACT Composite score.

The logistic regression model for ACT used in figure 3 used only ACT Composite score. To more thoroughly examine the predictive usefulness of ACT, another logistic regression model was developed to predict student on-time Associate degree completion, this time using all available ACT sublevel scores in addition to ACT Composite.

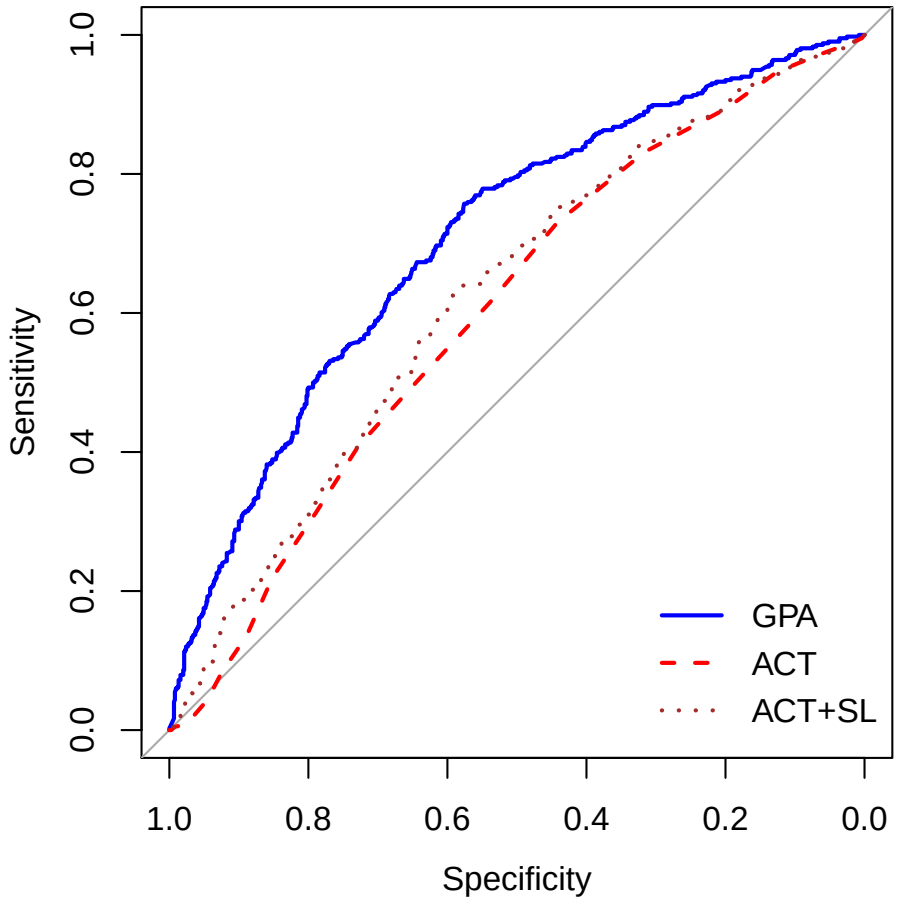

Figure 4: GPA vs. ACT Composite vs. ACT Composite+Sublevels ROC Curves

Not surprisingly, adding in sublevel scores increases the accuracy of the ACT Composite logistic regression model. The AUC for the ACT+Sublevel model is 0.624, compared to 0.602 for the model using ACT Composite only. With a bootstrap test p-value of $p = 0.0178$, we see that there is a statistically significant improvement in model accuracy when using sublevels in addition to ACT Composite score alone.

However, this improvement still leads to a model that is significantly different from the GPA only model. The AUC for the GPA model is still 0.706, and after conducting a significance test for a potential difference in the AUC between the GPA model and the ACT+Sublevel model, we find a p-value of $p < 0.0001$. So while adding sublevels increased the predictive accuracy of the ACT model, it is still statistically significantly less accurate than the GPA only model.

While, to this point, it is fairly clear that a student's high school GPA is more valuable in predicting on-time Associate degree completion than ACT information when squared off head-to-head, another model was constructed to determine if they are more useful together than simply GPA alone.

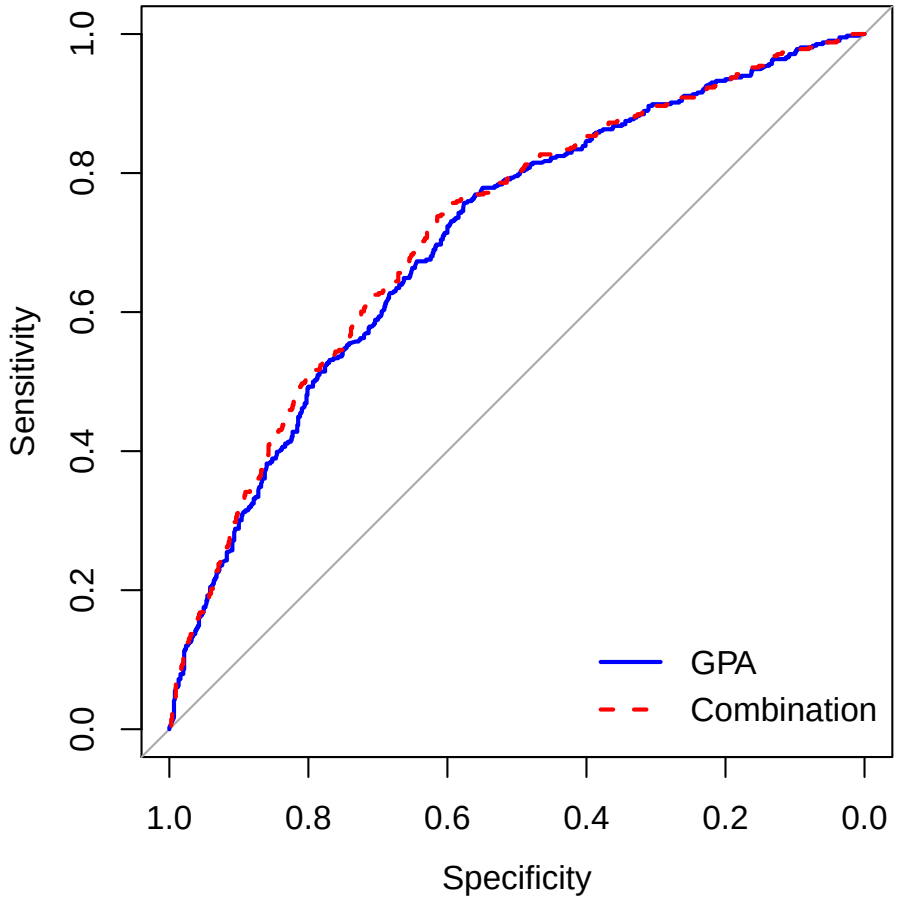

Figure 5: GPA vs. Combination Model ROC Curves

The combination of GPA and ACT Composite score does indeed result in a more accurate prediction model, with an AUC of 0.716, compared to the AUC of the GPA only model, 0.706. The test to determine if there was a statistically significant difference between the two AUC yielded a p-value of $p = 0.0061$. This indicates that, while there is not a large increase in AUC by adding ACT Composite to the GPA model, it is a statistically significant increase.

Summary

The goal of this study was to determine whether a student's high school GPA, ACT Composite scores, both, or neither were useful in projecting a student's ability to complete an Associate Degree on-time. By limiting the study to only students who have not transferred to an out-of-state university or an NDUS 4-year institution, this study is largely limited to students whose goal was to complete an Associate degree, rather than transfer credits to another university.

If the question being asked is "Which one is *more* useful?", the data clearly show that a student's high school GPA is more indicative, when compared to ACT scores, of a student's likelihood to complete an Associate degree within 3 years of initial enrollment. However, this does not mean ACT data is not at all useful for this purpose, as combining a student's ACT Composite score with their high school GPA resulted in prediction model that was more accurate than using a student's high school GPA alone.