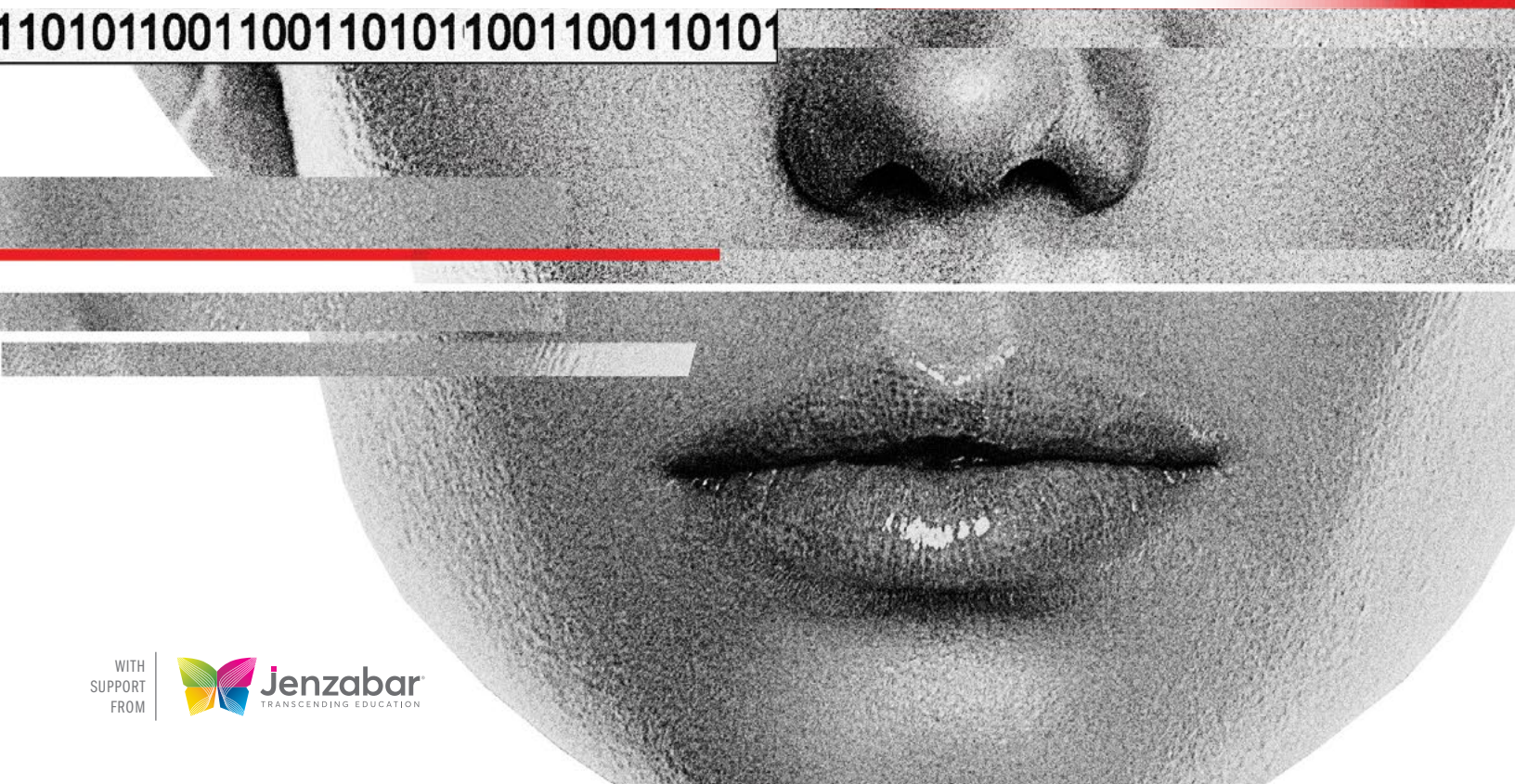
THE CHRONICLE OF HIGHER EDUCATION®

The Evolving State of Predictive Analytics

How it benefits students—

and where the tools fall behind



WITH SUPPORT FROM Jenzabar TRANSCENDING EDUCATION

Make informed decisions and take action.

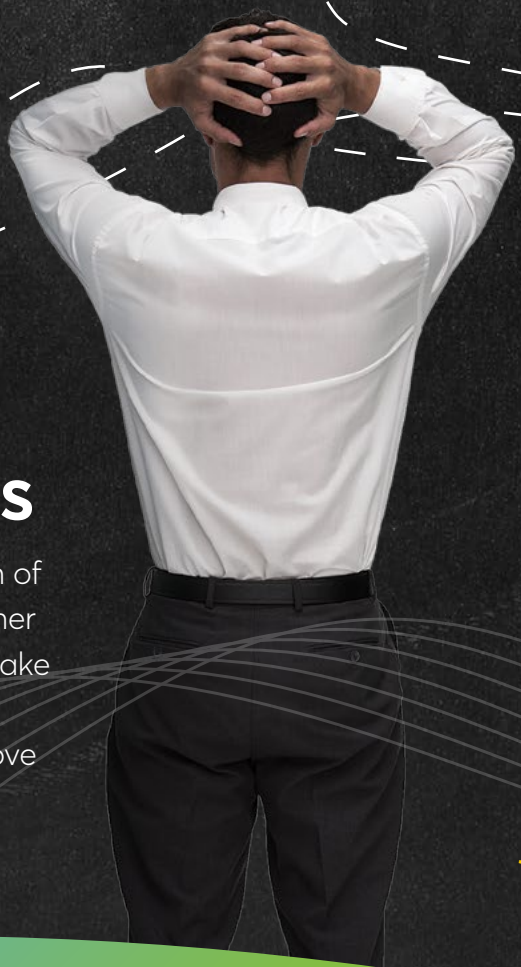
Jenzabar Analytics enables higher education leaders to use data to make smarter decisions, improve student outcomes, and achieve institutional objectives.

Academic Programs
How can we improve academic program performance?

Student Success
How do we get more students to completion?

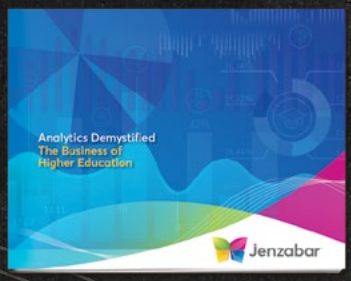
Financial Health
Which initiatives are supporting our long-term viability?

Student Enrollment
How many students do we need to enroll next year?



Jenzabar Analytics

Jenzabar Analytics is a platform of analytics tools that enables higher education campus leaders to make decisions and take appropriate actions to increase agility, improve performance, and identify new avenues for success.



Download our ebook to learn more about using analytics to make data-informed decisions.
jenzabar.com/ebook-CR

Jenzabar
TRANSCENDING EDUCATION

TABLE OF CONTENTS

The Evolving State of Predictive Analytics

Cover illustration by Lincoln Agnew

Contact CI@chronicle.com with questions or comments.

EXECUTIVE SUMMARY

For years, businesses have relied on collecting information and running it through algorithms to target customers with a range of offers from products to television series. Higher education has trailed in its use of predictive analytics, but rare is the college leader today who hasn't heard of the technique as a way to improve enrollment, retention, and graduation rates.

But how is predictive analytics typically used? What barriers exist for colleges that want to adapt the practice but have yet to do so? What are the successes and failures? To learn more, *The Chronicle* commissioned a survey of senior administrators and other campus officials.

The majority of respondents say their efforts to improve retention and graduation rates entail some use of predictive analytics — defined as "using historical data and quantitative techniques to help predict the probability of future events like enrollment and retention and inform interventions to encourage positive outcomes."

Most college leaders are confident that predictive analytics could improve student-retention rates, and that the practice has helped improve those rates on their own campuses. But only a minority of them use it campuswide, and some report problems, such as wrongly flagging students or groups of students as being at risk.

The reality is that predictive analytics cannot be viewed in a vacuum. As those who use and study it warn, it is just one tool.

> **Too often colleges invest in predictive analytics without understanding what it can and can't do.**

Predictive analytics has certainly been shown to help improve colleges' efforts to keep students from dropping out and get them to the graduation stage in four to six years. Especially today, during the coronavirus pandemic, that is more important than ever.

But too often, say those interviewed for this report, colleges invest in predictive analytics without understanding what it can and can't do. Too many expect the system to work without the need for good institutional data, fully training faculty and staff members, revamping programs

MAX WHITTAKER

California State U. at Sacramento tracks student phone data on campus in order to determine generalized patterns of behavior — which can then be used to better support at-risk students.

and policies, and understanding that the process must be constantly reviewed and updated.

To make colleges' use of analytics clearer, this report uses findings from the survey and interviews with campus leaders and other experts. Commissioned by *The Chronicle of Higher Education* with support from Jenzabar, the survey was conducted by Maguire Associates in March 2020. The research comprises results from 589 institutions; 87 percent are four-year institutions and the rest two-year or less than two-year. Not all questions were answered by all respondents.

INTRODUCTION

The hard lesson learned by most colleges when venturing into predictive analytics is that it's like buying an expensive gym membership — it's effective only if you put in a lot of effort.

"It's very complicated work that has taken a lot of institutions a long time to be successful at — and in the end, it's not about the data, it's the user," says Bridget Burns, executive director of the University Innovation Alliance, a consortium of 11 large public universities that are exploring ways — including predictive analytics — to increase graduation rates and reduce achievement gaps. "It's all about how it's executed and how you can actually drive a team to integrate and leverage the data."

The *Chronicle* survey reflects this complexity. Sixty-three percent of respondents say they use predictive analytics to improve retention and graduation rates, while 20 percent say they don't, and 17 percent aren't sure.

Almost half (46 percent) of the institutions that use predictive analytics say they have developed successful interventions, such as contacting students earlier about a missing assignment or reaching out to those who show signs of being at risk.

But only 20 percent use predictive analytics institutionwide, and a quarter say the tool has produced bad data or incorrectly flagged a student or group of students at their college. Fifty percent weren't sure if that had happened.

The survey "does align with my perception that many more institutions are purchasing or developing predictive-analytics tools to use in that context," says Martin Kurzweil, director of the Educational Transformation Program at Ithaka S+R, a consultancy and research organization. "It also is consistent with my findings that a lot of institutions that are purchasing those tools are having trouble working with them — the implementation is often quite difficult, and in a lot of cases they don't see the benefits that they expected."

RYAN FLOOD

Bridget Burns is the executive director of the University Innovation Alliance, which explores ways — including predictive analytics — to increase graduation rates and reduce achievement gaps.

Colleges have several ways to create a predictive-analytics system: outsource to a vendor, create their own systems from scratch, or use a combination of the two.

Typically, when a predictive-analytics system is used specifically for student success, it starts with an early alert, which flags a student who might be at risk on the basis of factors built into an algorithm. Those factors change as the student progresses through college; for example, a high-school GPA is important for a freshman, less so for a senior. Other factors include cumulative credits and completion ratios, which measure credit hours earned to credit hours attempted.

Whenever the term "predictive analytics" comes up, talk inevitably turns to Georgia State University and its success in closing its achievement gap and raising its six-year graduation rates from 32 percent in 2003 to 55 percent in 2018 — in a student population more than 60 percent nonwhite and one-third first-generation.

The university, which is a member of the University Innovation Alliance, is an exemplar of how to use predictive analytics; its advisers track more than 800 risk factors daily. It has also been highly effective at using forecasting beyond individual students to assess courses, programs, and policies that are limiting success, Kurzweil says.

But people are taking away the wrong lesson if they think predictive analytics is a static tool, Burns says. "Georgia State did not set it and forget it. They did not adopt predictive analytics, and then everything was fixed. They daily look at the data they're gathering and daily question whether these are the right indicators to focus on. They daily cross-apply those data indicators with what they're seeing in real-time interactions with advisers. This is something they have to consistently work at."

The University of Maryland-Baltimore County understands how difficult it can be to properly put in place the more widespread use of predictive analytics. It works with a vendor that uses an early-alert system to identify students who are in danger of failing by week seven of the semester. But that seems too late, says Robert Carpenter, associate provost for analytics and deputy chief information officer.

He is creating an in-house model with a team of undergraduate and graduate students that will provide information to instructors and advisers around the fourth week. "We're kind of cautious with our use of predictive analytics," he says. "The principle we're operating under is to do no harm."

DAVID GOLDMAN, AP IMAGES

A student walks by the library at Georgia State U., which has successfully closed its achievement gap and raised its six-year graduation rates thanks to careful use of predictive analytics.

SECTION 1

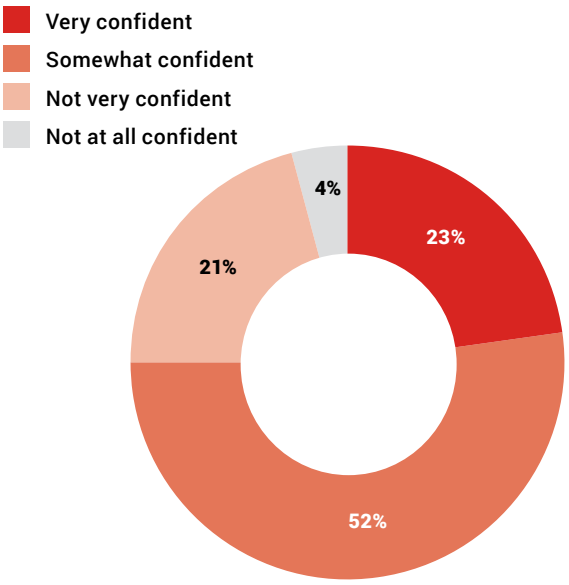
Obstacles to Use

Eighty-seven percent of those surveyed by *The Chronicle* believe that predictive analytics can improve retention and graduation rates at colleges around the country. But far fewer respondents say their college actually uses the tool.

Cost is an obvious barrier but certainly not the only one. Collecting the right data, determining clear goals, and properly training staff members in how to use the information all rank at the top when respondents were asked about barriers to using predictive analytics.

It's not surprising that the price tag is seen as a major impediment. Predictive-analytics systems, which are usually included in student-retention or student-success plans that do much more than forecasting, typically cost about $300,000, says James Wiley, principal analyst for technology at Eduventures, a

How confident are you that the use of predictive analytics has improved student retention and graduation rates at your institution?



Source: *Chronicle* survey

higher-education research-and-analysis firm now owned by the ACT. "And that's just for the tool," he says. "There's possible additional cost for integration and consulting services."

Wiley estimates that about 25 vendors offer predictive analytics, with five major players dominating the market. Vendors often assure colleges that the cost will be offset by increases in tuition revenue through higher enrollment and retention. But in the *Chronicle*'s survey, only 13 percent of 272 respondents say that's true for their own institution, with 42 percent disagreeing and 13 percent saying they're not sure.

John Smail, associate provost for undergraduate education at the University of North Carolina at Charlotte, puts himself in the third camp. His university's rates of graduation and retention have gone up significantly, he says, but "I can't tell you if it's paying for itself — that's not how our budget works. I know that's what companies like to say."

COURTESY OF UNC CHARLOTTE

John Smail, associate provost for undergraduate education at the U. of North Carolina at Charlotte, says he's not sure the cost of the tools can be offset by increased tuition revenue.

> **"A lot of times vendors walk in assuming that the institutions have a set of questions in mind, but in reality they don't."**

But cost is only the most obvious hurdle. Before hiring a vendor or developing an in-house system, colleges have to be very clear what they hope to get from predictive analytics — and that means figuring out "what questions do you really want to answer except just who's at risk," Wiley says. "You have to unpack that a little more. A lot of times vendors walk in assuming that the institutions have a set of questions in mind, but in reality they don't."

One rule of thumb, experts say, is that a college shouldn't do an analysis if it can't take action based on the data. Then there's the quality of the data. Many times "these third-party products fail," says Jay Golden, president of Wichita State University, "and it's not because the software engineers wrote bad software, but because most of time vendors cannot address the institutional data quality."

Gathering the information is not easy — and Golden has found this to be true not just at Wichita but at other institutions he has worked at. "You're reliant on staff and department chairs, all of whom have their own full-time jobs," he says. The sweat equity eventually pays off, but "academic, administrative, and research divisions operate on multiple and different platforms, and trying to get the different programs

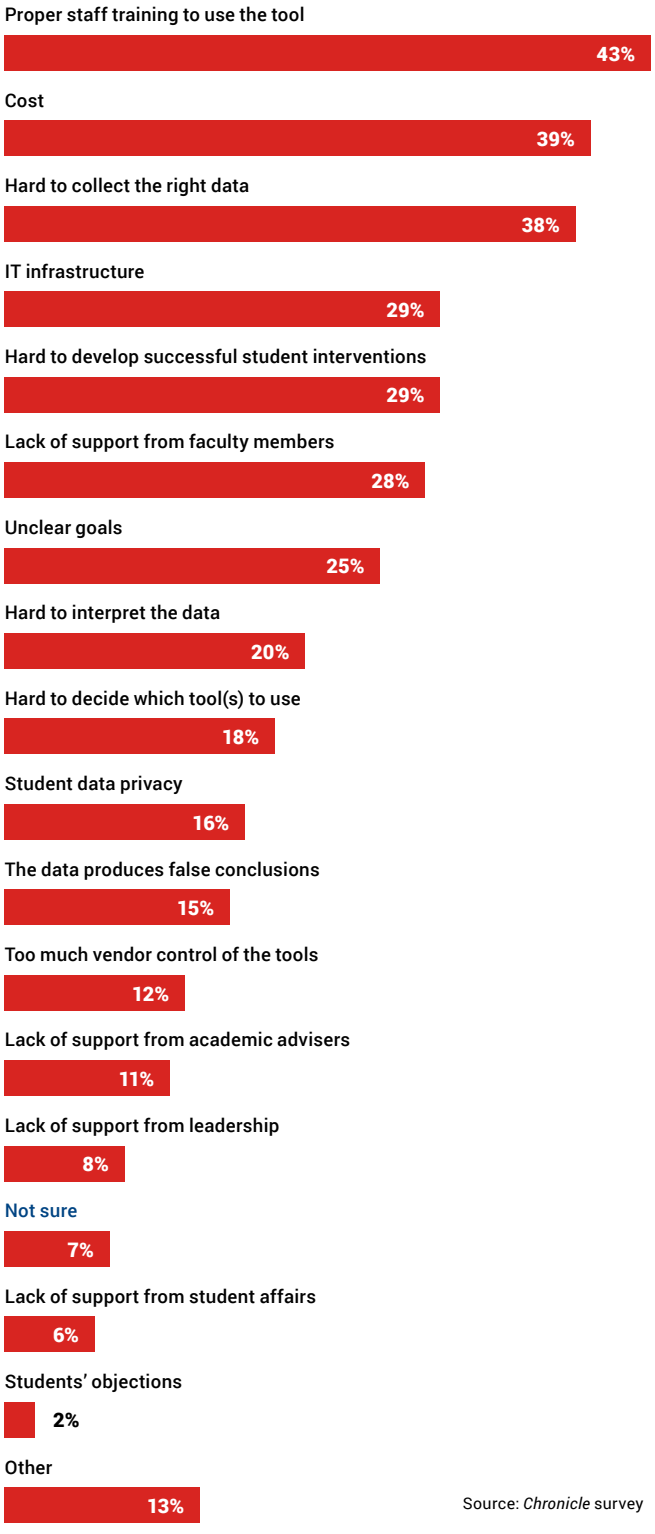
to talk to each other takes a lot of time — and I'm talking years."

While it may be simpler for colleges to buy predictive tools from a vendor, there is the black-box conundrum — the algorithms that companies use are rarely transparent, so institutions may not know what inputs are used.

Purdue University was one of the 11 institutions that joined the University Innovation Alliance's program on predictive analytics in 2015. "The original idea was that everybody was going to buy the same products and apply them the same way, and the magic would happen," says Frank Dooley, formerly senior vice provost for teaching and learning, and now chancellor of Purdue University Global. That's not quite how it worked. Each member of the alliance had its own successes and challenges, but an obvious obstacle early on was that data sets, student-information systems, and curricula were not consistent from university to university.

Although other parts of the student-success system worked well for Purdue, such as a notes platform for everyone who interacts with a student, "we looked at the predictive part really hard for almost two and a half years and finally said, 'We're not getting any value from it,'" Dooley says. Now Purdue is working on its own data analysis. "It's more us dictating the data than relying on a black box to tell us what to do."

What are the main barriers to using predictive analytics more widely at your institution? Select all that apply.



Source: *Chronicle* survey

SECTION 2

Tailoring the Process

At UNC Charlotte, advisers didn't like their vendor's risk scores of students. The flagging didn't always turn out to be wrong, but it was off just enough to make them doubt its efficacy. And many students at Charlotte are transfers, which created additional uncertainty about the validity of the risk scores, says John Smail, the associate provost for undergraduate education.

"We wouldn't ask advisers to reach out to all students who are flagged red," he says. "I couldn't tell you the percentage that's inaccurate, but the further a student goes into a program, the more accurate it is. At 60 hours it's better than 30 hours, but with incoming students, it's a flip of the coin."

The university has found, howev-

What key questions would you like predictive analytics to answer about the students at your institution?

- **"Which students can we successfully identify who may run into academic difficulty and what strategies should we put in place to support them?"**
- **"What array of programs will most effectively meet the needs of our students and the state?"**
- **"What drives students to enroll?"**
- **"How can we improve four-year graduation rates?"**
- **"The impact of demographics on predictions of success. What impact does the intersection of race and socio-economic status have on success?"**
- **"Does student engagement in student organizations and nonacademic functions have any correlation to academic achievement?"**

Source: *Chronicle* survey

er, that the information and the alerts can be useful when more finely tuned. Predictive analytics can help show which specific courses are important to ensure success in a major, and even what grade is needed — for example, in some courses, it would be worrisome if a student earned less than a B, while in others, a C would be fine.

"The red, yellow, green alerts weren't terribly helpful, but the more-nuanced data in the tool was helpful," Smail says, noting that every year "it's gotten a little more sophisticated."

In the *Chronicle* survey, 65 percent of the respondents say the use of predictive analytics affected institutional practices, with changes in academic advising being the most common. That is true at UNC Charlotte. Among other things, the university requires advisers to use the same platform, and it has built an advising culture that's more uniform across campus, so that if a student changes majors, the advising experience is more or less the same.

Seventy-five percent of the respondents also express confidence that predictive analytics has improved retention and graduation rates at their institution. Smail would agree. His university has seen an increase in its six-year graduation rate from 53 percent in 2015 to 64 percent in 2019 and in its retention rate from 77 per-

LINCOLN AGNEW

cent to 83 percent over the same period.

"I couldn't quantify it and couldn't prove it, but I think it's important," he says. "We know students do respond to predictive analytics-based interventions." In 2019 the Association of Public and Land-Grant Universities gave Charlotte its Degree Completion award.

Texas Tech University, using what it had learned from predictive analytics, changed advising in a number of ways. It created a coordinated-care team, made up of student-success professionals from every undergraduate college, to serve students at immediate and severe academic risk. A few years ago, it established a system of academic-life coaches, who are a stable presence in students' lives even if they change majors.

Since 2011, notes Patrick Hughes, vice provost for university programs and student success, Texas Tech's one-year retention rate for first-year, first-time students, has increased from 81 percent to almost 87 percent. But "some of what we learned — and one key aspect that often gets overlooked at universities embarking on a student-success program that includes data analytics — is how to get faculty involved."

That means, among other things, working with professors on how they want to report on student risk. Predictive analytics can take into account static data, such as GPA and demographics, and generate a risk assessment, but what can be lacking is feedback from faculty members that contributes to the overall risk model. That's important when you're trying not just to identify students who may be in danger of failing but also to figure out the best intervention to mitigate that risk.

There's another issue: Predictive analytics can do a good job of alerting an adviser to reach out to a student. But getting that student to respond is not so easy. Smail would like to see more work figuring out how to make that happen.

"We call it the 'getting students' attention problem,'" he says. "We're getting better, but our data suggests that this is actionable information that students are not responding to." Only about 40 percent of students flagged as emerging risks who hear from an adviser get back to that adviser.

> **"There's great efficacy to having good data. But the data do not retain students. It's what we do with our data and how we communicate to students based on that data that makes the difference."**

Charlotte found that students who, by the fourth week of the semester, had been flagged with an early alert in at least two classes, with one of the classes being important to their major, and who responded to outreach from their adviser, had statistically significant improvement in their GPAs. They also earned more credits in their courses than those who did not respond to an adviser.

"There's great efficacy to having good data," Smail says. "But the data do not retain students. It's what we do with our data and how we communicate to students based on that data that makes the difference."

Learning how to improve student response, he adds, is "the challenge of the next couple of years."

SECTION 3

Privacy and Nudges

The use of predictive analytics raises concerns aside from efficacy. The potential to gather data to make the analytics increasingly precise — drilling deeper into students' lives and seeing how often they visit a gym, say, or a library — makes some people nervous. A particular worry is that much of this effort will be focused on low-income students and students of color, who also tend to make up the larger at-risk student populations.

"The concern with a lot of surveillance in higher education and in general society is

LINCOLN AGNEW

> **"The concern with a lot of surveillance in higher education and in general society is that populations that bear the brunt tend to be disenfranchised and unprivileged."**

that populations that bear the brunt tend to be disenfranchised and unprivileged," says Kyle M.L. Jones, an assistant professor of library and information science and data science at Indiana University-Purdue University at Indianapolis. "I understand using predictive analytics for this population — we're trying to give information and resources to help them stay in school and be successful. Yet when you compound with government and commercial surveillance, where does it stop?"

The answer, Jones says, is for colleges to be more transparent about what they're doing. "We need to be more forthcoming to our students about what data we have, how we use it, who has access to it, and who doesn't," he says, adding that a few colleges are drawing up policies accordingly, following the example of the Open University, in Britain.

Another fear is whether too many students who may be struggling early on in a course for one major might be nudged into a major in which completion is more likely. That did not occur at Georgia State University, says Bridget Burns, of the University Innovation Alliance. When students there changed majors, they tended to move toward more-difficult ones. "They're guiding students to make good choices, not driving them toward an easier major," she says.

But as Burns and others acknowledge, every college will use predictive analytics in different ways, "I worry about that," says Purdue's Frank Dooley, "especially for low-income and first-generation students. There's the danger of having technology think for you."

As Robert Carpenter, of UMBC says, "we want to use predictive analytics to hold doors open for people, not to close them."

Please briefly describe an example of how the use of predictive analytics has improved a practice on your campus

- **"Establishment of a number of co-req courses to help fill gaps in student learning without delaying progress toward a degree."**
- **"A change in policy to the add/drop deadline where withdraw (W) is not punitive like a failing grade (D or F), but a success strategy to preserve GPA."**
- **"Changing financial-aid packaging formulas; identifying tactics that might have a positive impact on student retention."**
- **"Using analytics, the registrar is able to help associate deans streamline the scheduling practices of departments to better align schedules with student needs."**
- **"We are changing prospective student event dates based on analysis that students are applying but not yielding earlier."**
- **"We now will prevent students from registering in multiple high failure rate courses in the same semester now that they have been identified."**

Source: *Chronicle* survey

SECTION 4

What's Ahead

If there was ever a time that predictive analytics might fail, it's during an unprecedented event such as a pandemic. As almost all colleges moved to remote learning in March, the usefulness of any data gathered remains an open question.

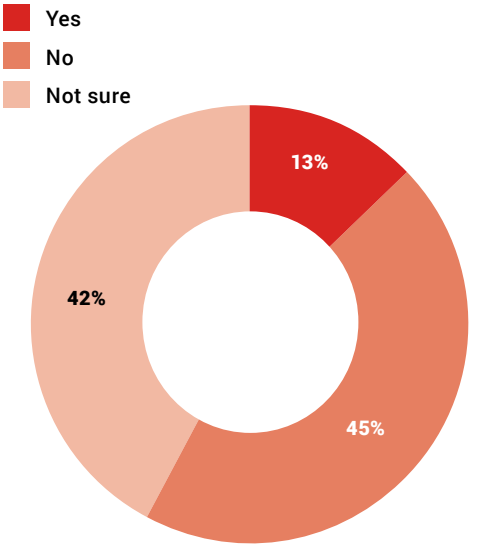
In some ways, because students are all online now, there's an opportunity to generate more data points, contributing to more-sophisticated algorithms. Colleges that already have tools for student success in place on the basis of predictive analytics "will be in a really good position to help students who might be struggling right now," says D. Christopher Brooks, director of research at the Educause Center for Analysis and Research. "For example, early-alert systems can identify students who might not be logging into their courses, participating in class, handing in assignments, or downloading materials."

Martin Kurzweil, of Ithaka S+R, says it's important not to abandon the information now coming out of the systems, but to treat it more skeptically.

Beyond the current crisis, however, campus leaders who have immersed themselves in the use of predictive analytics to raise graduation and retention rates have thoughts on where they would like to go next. They want to look more deeply at curricula and analyze their complexity, both in individual courses and in relation to one another, as well as find ways to avoid student-information fatigue, so advisers don't become desensitized to data and the need for interventions.

"We have to think of student success not in overnight terms but maybe in generational terms," says Patrick Hughes, of Texas Tech, "as these platforms become more user-friendly and more insightful with each generation."

Has the cost of implementing predictive analytics been offset by increases to tuition revenue from increased enrollment or retention?



Source: *Chronicle* survey

THE CHRONICLE
OF HIGHER EDUCATION®

1255 Twenty-Third Street, N.W.
Washington, D.C. 20037

(202) 466-1000 | Chronicle.com