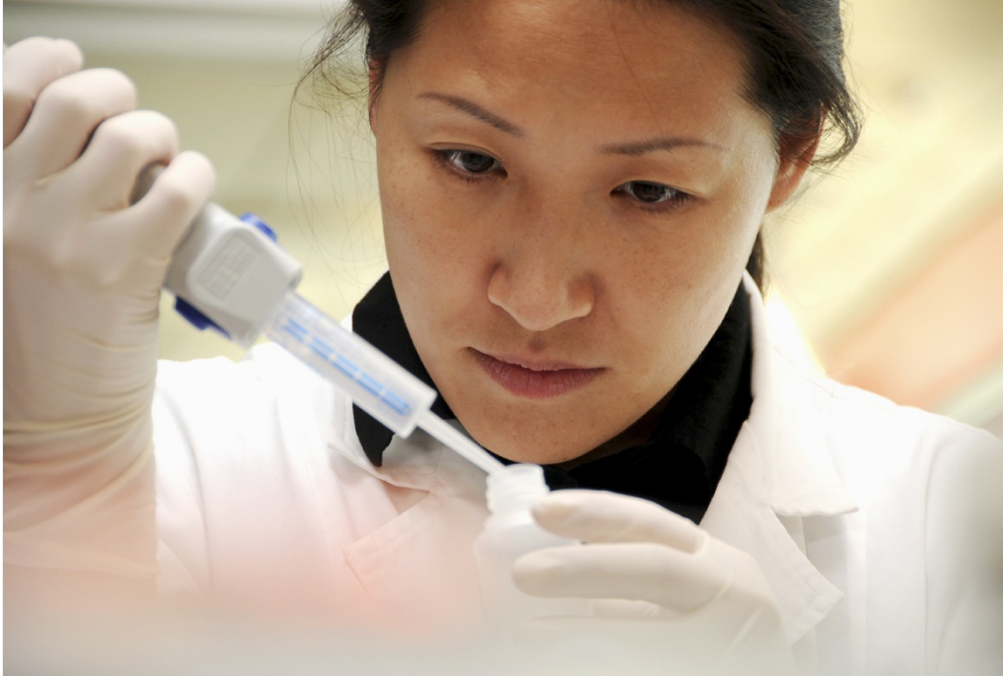


# Computer model predicts academic success

Algorithm based on publications finds that first-author articles in leading journals matter most.

Richard Van Noorden

02 June 2014



Norman Jean Gargasz/Getty Images

Female biologists are less likely to become principal investigators than are male biologists with comparable publication records, a statistical model found.

The mantra 'publish or perish' is drilled into every early-career scientist — and for good reason, a computer model suggests. The most important predictor of success for a young biomedical scientist is the number of first-author papers published in journals with high impact factors early in a researcher's career, according to the formula.

The model, created by computer scientist Lucas Carey, at Pompeu Fabra University in Barcelona, and his collaborators, also found that, even correcting for publication records, working at a highly ranked university — and being male — are predictors of academic success.

The results appear today in *Current Biology*<sup>1</sup>, and the team has built a website, [pipredictor.com](http://pipredictor.com), which tells early-career scientists whether their publication records are ahead of the norm.

Together with David van Dijk at the Weizmann Institute of Science in Rehovot, Israel, and Ohad Manor at the University of Washington in Seattle, Carey analysed the publication records of more than 25,000 biomedical researchers who first published between 1996 and 2000. Just 6.2% of these scientists would go on to become a principal investigator (PI), which the team defined as publishing at least three papers as the last-named author, after 13 years. Their algorithm used machine-learning techniques to analyse which of more than 200 different metrics were the best predictors of success, and produced a formula that can be used to compare the odds of two individuals from the same cohort becoming a PI. The formula picks out the more successful researchers with an 83% success rate.

## Middle-author obscurity

According to the formula, the quantity of publications in a scientist's first eight years, and the impact factors of the journals in which they appear, are highly influential factors. But it is the first-author publications that count most, adds Carey. "It's really bad to be a middle author," he says.

However, first authors who publish with a large number of co-authors start to lose their advantage, Carey explains. "Our data suggest

you shouldn't get trapped in big projects," he adds. And other measures of quality — such as the citations a paper receives above the norm of the journal — fade beside the importance of the journal impact factor, even though many scientists castigate this metric as a flawed way of judging the quality of researchers.

Being male is also a positive predictor for becoming a PI, the results suggest. On average, having an identical publication record but being a woman lowers the chance of success by 7%, for example, says van Dijk. Those at leading universities also had much higher chances of becoming a PI — although when this effect was untangled from good publication records, university rank on its own was not particularly influential.

Some scientists defied the predictions, becoming PIs even though they never published in high-impact journals. But those authors had twice as many first-author papers as their competitors who didn't become PIs — suggesting that enough first-author papers can compensate, the team found.

"The paper shows that academic success has clear predictable features, and decomposes cleverly the main factors that contribute to it. In many ways, it really indicates the powerful role of appropriate metrics in gauging academic success," comments Albert-László Barabási, a network theorist at Northeastern University in Boston, Massachusetts, who has published work on predicting academic success at the level of individual papers.

"Most of the findings line up reasonably well with career advice that young researchers are given all the time," says Zen Faulkes, a biologist at the University of Texas–Pan American in Edinburg.

But the method has its limits. The formula has been tested only on data about biomedical scientists in the 1996–2000 cohort, so the factors for success might be different today. And the model's apparently high predictive accuracy may be misleading, notes Peter van den Besselaar, a researcher who studies the management of science at the Free University of Amsterdam. It is relatively easy to accurately distinguish between the low-performers who will leave academia from the high-performers who are likely to become PIs, he says; the hard part is divining who will be successful among just the top 15% of scientists (of whom fewer than half will become PIs). And the model's definition of a PI does not necessarily line up with what people might think of as academic success, such as becoming head of a lab or gaining tenure.

Still, the analysis is useful, Faulkes adds, because it documents the disadvantage of female researchers, and shows that the journal impact factor still holds a lot of power.

Carey says that the team is now working on computer models that try to predict who might get tenure, and who will be awarded a grant. "From my experience, what university hiring committees care about is mainly how many papers you have and whether you have funding — and all the funding agencies care about is your publication record," he says.

*Nature* | doi:10.1038/nature.2014.15337

## References

---

1. Van Dijk, D., Manor, O. & Carey, L. B. *Curr. Biol.* <http://dx.doi.org/10.1016/j.cub.2014.04.039> (2014).