

## Commentary

# The case for caution in predicting scientists' future impact

Any scientist pursuing a research career these days is acutely aware of the increasingly central role metrics play in measuring scientific impact. From papers to people, the quality of almost everything is being measured by citations. Publication metrics are starting to shift the way in which scientists are having their career potential evaluated. From an economic point of view, a tenure-track hire is a million-dollar bet on a young scientist's future success, so it is easy to see why predictive metrics and models are attractive to decision makers. Million-dollar gambles show that the genie of metrics and models is unlikely to be put back in the bottle.

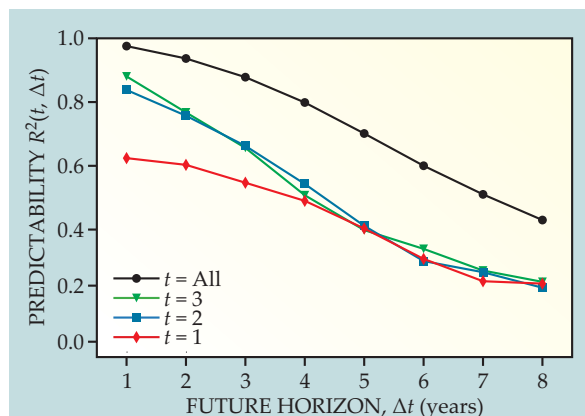
If metrics are to be integrated into the career advancement process, they must be better tested, and specific issues must be addressed: What aspects of a career are predictable? What ingredients make a model robust? How often is a given model's prediction wrong, and what impact do errors have on the careers of scientists, especially young ones already burdened by risk? Without clear answers to these and other questions, the unexamined use of quantitative indicators can do real harm not only to scientists who may be shown the door based on bogus evaluations but to the endeavor of science as a whole.

Jorge Hirsch's introduction of the  $h$ -index<sup>2</sup> in late 2005 was a significant milestone in the use of metrics in career evaluation. The popularity of the  $h$ -index has been steadily growing since its introduction. In fact, it now stands as the most popular quantitative measure

of a researcher's productivity and impact. It is already being used to evaluate scientists; a modified version has been integrated into the Italian national tenure competition overseen by the National Agency for the Evaluation of Universities and Research Institutes.

Future impact, rather than past accomplishment, is at the heart of most science-career appraisal decisions regarding tenure, grants, fellowships, prizes, and so forth. Hirsch's additional work indicates that the  $h$ -index is better than other indicators in predicting future scientific achievements.<sup>3</sup> A more recent publication by Daniel Acuna and coworkers presents a model that uses a linear combination of five metrics to predict an individual's future  $h$ -index.<sup>4</sup> The technical details of that work are notable because it is one of the first models to integrate several metrics into a prediction. However, some of its nontechnical aspects are probably more noteworthy: It was published in a high-profile forum; the authors suggest that it can be used in decision making; and it even includes an online future  $h$ -index calculator.

In the model from Acuna and coworkers, a future  $h$ -index,  $h(t + \Delta t)$ , is calculated from a linear combination of five metrics: an individual's current  $h$ -index  $h(t)$ , the square root of his or her total number of publications  $N$ , the number of years  $t$  since first publication (the career age), the number of publications  $q$  in high-impact journals, and the number of distinct journals  $j$  in which the individual has published. With its use of several key metrics of academic publishing, the Acuna team's multiple regression model appears quite promising. However, further investigation highlights the care that must be taken in developing models of future impact.



**A measure of future success?** Although the Acuna model,<sup>4</sup> based on five metrics of a scientist's publication output, is respectably predictive when all age cohorts in a data set of 100 US assistant professors in physics are combined ( $T = \text{All}$ , black curve), it decreases significantly when early-career-age cohorts whose years since first publication  $t = 1, 2$ , or  $3$  (red, blue, and green curves) are analyzed separately.

To illustrate the difficulties of predicting future success, we applied the Acuna model to a career data set of 100 assistant professors in physics, two from each of the top 50 physics departments in the US (see reference 1 for a further description of the data set). The figure above shows the coefficient of determination  $R^2(t, \Delta t)$ , a statistical measure of how well the model predicts,  $\Delta t$  years into the future, the  $h$ -index of a scientist with academic age  $t$ . The Acuna model aggregates all years in the data sample ( $t = \text{All}$ , black curve), and in doing so it yields a respectable prediction of  $h(t + \Delta t)$  even up to  $\Delta t = 6$  years. However, we find that the model's predictive power arises largely *because* all the career-age cohorts are combined.

To demonstrate the point, we also show the Acuna model  $R^2(t, \Delta t)$  calculated using separate early-career cohorts  $t$  (green, blue, and red curves). The  $R^2(t, \Delta t)$  values calculated for a fixed  $t$  are significantly less than those calculated by aggregating across all career ages: The model is generally poor at predicting the future success of early-career scientists. The limitation is of particular concern because early-career decisions

Letters and commentary are encouraged and should be sent by email to [ptletters@aip.org](mailto:ptletters@aip.org) (using your surname as the Subject line), or by standard mail to Letters, PHYSICS TODAY, American Center for Physics, One Physics Ellipse, College Park, MD 20740-3842. Please include your name, work affiliation, mailing address, email address, and daytime phone number on your letter and attachments. You can also contact us online at <http://contact.physicstoday.org>. We reserve the right to edit submissions.

make up a significant portion of cases in which quantitative approaches are likely to be applied. By using additional career data for 200 highly cited physicists, we further confirmed our observation of much lower  $R^2$  values in the early career ( $t$  up to 3 or 4 years).

Recent work by Amin Mazloumian hints at one of the underlying difficulties of predicting a scientist's future success.<sup>5</sup> By differentiating between citations accrued by papers already published at the time of prediction and citations accrued by papers published after the prediction time, Mazloumian shows that regression approaches do a reasonable job of predicting future citations to past papers but do not reliably predict future citations to future papers. Therefore, those who would predict a scientist's future value should be aware that the impact of papers published in the past does not necessarily correlate with that of papers published in the future.

Going forward, the metric-based approaches and their successors will be increasingly exploited in decision-making processes. However, little is known presently about the strengths and weaknesses of the state-of-the-art predictive indicators. Where the responsibility lies for vetting current and new quantitative measures is still an open question. But scientists themselves, particularly young ones, clearly stand to lose the most should quantitative measures be weighted too heavily in decisions affecting their careers. It behooves us to engage with the institutions that seek to use these quantitative measures of impact in their decision making and to impress upon them a skepticism backed up by quantitative and rigorous analysis of the specific measures they seek to employ.

## References

1. A. M. Petersen, M. Riccaboni, H. E. Stanley, F. Pammolli, *Proc. Natl. Acad. Sci. USA* **109**, 5213 (2012).
2. J. E. Hirsch, *Proc. Natl. Acad. Sci. USA* **102**, 16569 (2005).
3. J. E. Hirsch, *Proc. Natl. Acad. Sci. USA* **104**, 19193 (2007).
4. D. E. Acuna, S. Allesina, K. P. Kording, *Nature* **489**, 201 (2012).
5. A. Mazloumian, *PLoS ONE* **7**(11), e49246 (2012).

Orion Penner

Alexander M. Petersen

IMT Institute for Advanced Studies Lucca  
Lucca, Italy

Raj K. Pan

Santo Fortunato

Aalto University School of Science  
Aalto, Finland

## Letters

### Strange connections to strange metals

The June 2012 issue of PHYSICS TODAY has, beginning on page 68, the Quick Study "From black holes to strange metals," by Hong Liu. It is one of many quasi-journalistic discussions I have seen of results using the AdS/CFT (anti-de Sitter/conformal field theory) correspondence from quantum gravitation theory ostensibly to solve condensed-matter physics problems such as the "strange metal" in the cuprate (high  $T_c$ ) superconducting metals. As the probable source of the buzzword phrase "strange metal" to describe the phenomena observed in the cuprates and of a theory that bids well to explain those phenomena in detail, I think I have a reasonable motivation to object to the publication of those claims, even though advanced tentatively, when so much is known about this particular phase.

The strange-metal region of the cuprate phase diagram exhibits not only a linear dependence on temperature of the conductivity relaxation rate, which is generally taken by string theorists as *the* characteristic symptom identifying a strange metal and is the only feature they discuss. The region also exhibits several additional anomalies that in my experience are unique to this phase:

► The IR conductivity—the "Drude tail" of the DC conductivity—falls off with frequency with a noninteger power law, and the exponent apparently varies continuously with doping. That behavior was demonstrated by Nicole Bontemps and coworkers<sup>1</sup> in 1993 and further nailed down by Dirk van der Marel and coworkers<sup>2</sup> in 1995.

► The relaxation rate as measured by the Hall angle  $\theta_H$  of deviation of the current from the electric field direction, using the formula  $\theta_H = 1/\omega_c\tau$ , is quite different from that of the conductivity, and has a different,  $T^2$  temperature dependence, as N. P. Ong and coauthors demonstrated<sup>3</sup> in 1989.

► Over broad regions of doping, the two kinds of relaxation rates, the one for the conductivity and the one for the Hall rotation, seem to add as inverses: Conductivity is proportional to  $1/T + 1/T^2$ —that is, it obeys an anti-Matthiessen law.

► Angle-integrated photoelectron spectra, tunneling spectra, and angle-resolved photoemission spectra all



**SIM984 ... \$995** (U.S. List)

- Isolation up to  $\pm 1000$  V
- 1 MHz bandwidth
- 15 nV/ $\sqrt{\text{Hz}}$  input noise
- 1 M $\Omega$  input impedance
- Selectable gain & bandwidth

The SIM984 is a wide bandwidth, low-noise isolation amplifier with gain of up to  $\times 100$ . The input stage floats up to 1000 V above or below ground and is optically coupled to the output stage. The all analog design insures no quantization artifacts can corrupt sensitive signals.

In addition to isolating and amplifying signals, the SIM984 can also be used to break ground loops and reduce common-mode interference.



*SIM900 Mainframe loaded with a variety of SIM modules*