

THIS WEEK

EDITORIALS

EXOPLANETS The past, present and future of the hunt for extraterrestrials **p.272**

WORLD VIEW Global statistics need a new worldwide watchdog **p.273**



DEEP STORAGE Frozen turtles hibernate but stay sharp **p.274**

The maze of impact metrics

In deciding how to judge the impact of research, evaluators must take into account the effects of emphasizing particular measures — and be open about their methods.

So much science, so little time. Amid an ever-increasing mountain of research articles, data sets and other output, hard-pressed research funders and employers need shortcuts to identify and reward the work that matters. They have plenty of options: research impact is now recognized as a multidimensional affair.

The conventional measures of scholarly importance — citation metrics, publication in influential journals and the opinion of peers as expressed in letters and interviews — still loom large. But to those are now added metrics such as article downloads and views, and measures of importance beyond the academic realm, including influence on policy-makers or health and environment officials, effects on industry and the economy, and public outreach.

Researchers at the Center for the Study of Interdisciplinarity, part of the University of North Texas in Denton, this year came up with 56 measures of impact (see *Nature* 497, 439; 2013), including influence on curriculum creation, authorship of textbooks and success in surveys of colleagues' esteem. Some of these measures are a little fanciful, but they demonstrate that it has never been easier for scientists to show off the various ways in which their work deserves attention — and funds.

That variety is worth celebrating, but it can lead to dizzying confusion. How are researchers and evaluators to choose between measures? In this issue, *Nature* looks at some traditional and emerging ways to track research quality (see page 287). Ultimately, it is for institutions and funders to choose their preferences, but in doing so they should take two important considerations into account.

First, it is important to be aware of the positive and negative effects of privileging certain measures.

For example, emphasizing that research is considered especially important if it is published in one of a few historically influential journals — *Cell*, *Nature*, *Science* — could be a laudable attempt to get scientists to think ambitiously about their research goals. But it can also result in excessive pressure to publish big claims, leading to problems of irreproducibility, for example. (*Nature's* position is that it has been publishing research using essentially the same criteria for decades; it is up to the scientific community and evaluators to decide how much importance they want to place on papers that appear in the journal).

It is a mistake to consider a research paper important because it is published in a journal with a good citation record, as measured by its impact factor. As this publication has highlighted many times (see in particular *Nature* 435, 1003–1004; 2005), two articles in the same journal may have very different citation records. It is much better to focus on the citations, views or downloads of an individual article — and to recognize that these metrics vary between research disciplines.

In another example, emphasizing the economic impacts of research may force scientists to think about justifying their taxpayer-funded work, but it also runs the risk of distracting them with the lure of meaningless patents and ill-considered spin-out companies.

The second important consideration is the need for research evaluators to be explicit about the methods they use to measure impact. Openness is an essential part of earning trust. Evaluators should publish worked examples showing how they score assessments, and the reasoning behind such scores; even better would be, where possible, to publish the full data. Otherwise, researchers might rightfully feel suspicious (see, for example, writer Colin Macilwain's scepticism towards performance metrics: *Nature* 500, 255; 2013).

“It has never been easier for scientists to show off the various ways in which their work deserves attention — and funds.”

When scientists rail against the ‘impact agenda’, their arguments sometimes founder on irrelevant confusion between terms: too often, such discussion devolves into attacks on misuse of the impact factor, rather than looking at the range of possible metrics. The journal citation measure gains misleading prominence because its name happens to include the word impact — a semantic synergy that can cloud debate.

Arguments against impact metrics are strongest when they reference cases in which evaluators do not heed the considerations we mention above: in which evaluators choose metrics blindly, without sufficient thought for pernicious effects, or are secretive or inconsistent about their methodologies. If evaluators are to earn the acceptance — rather than the scorn — of the scientists whose work they want to fund, they had better pay attention to these concerns. ■

High hopes

Care must be taken not to raise unrealistic expectations for RTS,S malaria vaccine.

Vaccines have been an unparalleled public-health success: they have eradicated smallpox and driven polio to near extinction, and routine childhood immunization saves millions of children a year from death from diseases such as measles, diphtheria, tetanus and whooping cough. So it is not surprising that the public tend to view vaccines as synonymous with elimination, or near elimination, of our microbial foes.

This may help to explain last week's extensive and often upbeat media coverage of the 18-month results of a huge phase III trial of the malaria vaccine candidate RTS,S/AS01 in more than 15,000 children across 7 African countries. In the United Kingdom, for example, the front page of *The Guardian* stated that the vaccine “could save lives of millions of

children? Unfortunately, however, it won't. The 18-month results only confirm the disappointing results seen after 12 months.

The RTS,S vaccine is not what most people would think of as a vaccine. It provides only partial protection and most of those vaccinated, particularly those in areas with moderate to high malaria transmission rates, will eventually contract the disease. There is also confusion over its efficacy. Many media reports concluded that although the vaccine did not give the 90%-plus efficacy levels of most childhood vaccines, it might nonetheless be satisfactory, with a reported 46% reduction in cases in children vaccinated when they were aged 5 to 17 months, and 27% in 6–12-week-old babies.

Not so. The efficacy figures given for RTS,S are not directly comparable with those usually given for vaccines. The conventional measurement of a vaccine's success is how many people remain protected after a given period, such as 12 months. Because RTS,S is only partially protective, a different measurement of efficacy is used — a complex statistical model that computes hazard ratios on the basis of the first clinical episodes of malaria. As the designers of the method themselves concede, “a shortcoming of the vaccine efficacy calculated from hazard ratios could be that it is not intuitively understood”. Too true. In the hands of experts, and regulatory agencies, this hazards-ratio model offers a valid measurement of the efficacy of a partially protective vaccine, but it can be easily misinterpreted by the media, politicians and policy-makers.

It is not possible for outside scientists to deduce a more conventional efficacy estimate from the 18-month data, as it was described only briefly in press releases from the vaccine's sponsors, the PATH Malaria Vaccine Initiative (MVI) based in Seattle, Washington, and GlaxoSmithKline (GSK), headquartered in Brentford, UK. (The paper and supporting data are under review at a journal.) But applying a conventional measurement of vaccine success to the published figures for 12-month estimates — for which detailed data are available — reduces the vaccine's efficacy by more than one-third (see *Nature* 478, 439–440; 2011). Its protective effect also seems to begin fading after about six months.

Perhaps more promising are the reductions seen in cases of severe malaria, which are reported in the conventional manner. However, although a 36% reduction was reported in children of 5–17 months, the 15% reduction seen in 6–12-week-old babies was not significant — and this age group was the main target of the trial because for logistical reasons it is likely that any malaria vaccine would need to be given alongside routine immunizations at this age.

“The vaccine falls short of the target set by the World Health Organization.”

Many vaccine trial participants had access to other anti-malarial measures — including insecticide-treated bednets and effective drug treatment — so it is possible that the vaccine might offer greater benefit to people more exposed to malaria. Nonetheless, the vaccine falls short of the target for a partially protective malaria vaccine set in 2006 by the World Health Organization, which stated that it should have a “protective efficacy of more than 50% against severe disease and death” that “lasts longer than one year”.

The work will continue. Data on the effects of a booster dose given after 18 months will not be available until next year, and RTS,S is also due to be tested in combination with a vaccine developed by researchers at the University of Oxford, UK, in an early-stage clinical trial. Meanwhile, the RTS,S trials are to be applauded for having left a lasting legacy in the unprecedented collaboration with African scientists who led the study, and a first-class clinical-trials infrastructure on the continent.

RTS,S has been in the works for almost 30 years. Since 2001, the MVI has put some US\$200 million into it, and GSK more than \$350 million, with a further \$260 million earmarked to complete its development. The huge past impact of vaccines risks fuelling illusions over the impact of having a malaria ‘vaccine’. But the modest efficacy of RTS,S means that it falls squarely in competition with other malaria control measures, many of which might be more cost-effective. Care must be taken not to build excessive expectations that can only lead to disappointment over its potentially limited public-health impact. ■

Searching for life

A look into the past frames our attempts to find extraterrestrial intelligence.

Carl Sagan's 1993 *Nature* paper has, rather appropriately, a hint of science fiction about it. Twenty years ago this week, Sagan and a team of other astronomers announced that they had found life on a planet in the Galaxy. They used data from the Galileo space telescope to catch clear signatures of methane and carbon dioxide in the planet's atmosphere and abundant water in frozen and liquid states on its surface. They even confirmed the presence of radio emissions emanating from it — the canonical autograph of intelligence.

This month's *Nature* PastCast — one of a series of special audio treats available for free on the *Nature* website — recounts the tale. The twist, of course, was that this planet was Earth. Sagan and his team were trying out a method for finding life on other planets, using Earth as a calibration for future missions that might explore the depths of the Galaxy for signs of life.

Those were not friendly times for thinking about life elsewhere. At the time, the US Congress was debating whether to cut federal funding for NASA's SETI programme, the search for extraterrestrial intelligence. So Sagan and his team set about their task in as objective a way as they could, notwithstanding their foregone conclusion. They were careful to declare that “life is the hypothesis of last resort”, and to show that this was a scientific question that needed an answer.

The bigger question, of course, goes unanswered, although not for want of trying. SETI was launched in the late 1950s, propelled

by the optimism of the space age. In 1959, a paper in *Nature* by Giuseppe Cocconi and Philip Morrison suggested that if civilizations elsewhere wanted to contact Earthlings, they would probably use electromagnetic signals. “We shall assume that long ago they established a channel of communication that would one day become known to us, and that they look forward patiently to the answering signals from the Sun which would make known to them that a new society has entered the community of intelligence,” they wrote.

Soon after, astronomer Frank Drake was preparing for one of the first conferences to address the search for extraterrestrial life. As a loose agenda, he came up with a list of unknowns that would need to be resolved in order to predict whether intelligent life exists elsewhere in the Universe. For example, how many star systems exist that are suitable for the development of intelligent life? How many Earth-like planets are in orbit around them? What is the probability of life sparking into existence on any of them? Drake then formulated an equation that created a mathematical framework for such unknowns.

Research ongoing since Sagan's paper is making Drake's equation more solvable today than it has ever been. The control test was performed, so astronomers know that their tests for life would work. Meanwhile, the first exoplanet was found in 1992, and hundreds have been spotted since.

Scientists can use variants of Sagan's prescient control test to characterize the atmospheres and locations of exoplanets whizzing around their stars. Are we now in an era not of space-age optimism, but of realism? Life is still the hypothesis of last resort for astrobiologists. But if they find none, they will not be disillusioned. It would be just as interesting, they say, to find that habitable-looking environments do not all sprout life, and that Earth is unique in being so full of it. ■

➔ **NATURE.COM**
To comment online,
click on Editorials at:
go.nature.com/xhunjv