the Bush administration subsequently missed the legal deadline in 2004 to complete a second assessment, environmentalists sued the government in federal court to compel the report's release — and won.

The message of the latest science report — that human-caused global warming poses urgent problems for the United States — isn't likely to sit well with the White House. The Trump administration has sought to repeal environmental regulations and cut climate research. Energy secretary Rick Perry has joined Pruitt in questioning climate science. And Pruitt's chief of staff, Ryan Jackson, once worked for Senator James Inhofe (Republican, Oklahoma), a prominent climate sceptic.

"This is going to be the first big test in the climate arena," says Tammy Dickinson, who led the energy and environment division at the White House Office of Science and Technology Policy (OSTP) under president Barack Obama. One major issue, she adds, is that Trump has yet to fill many positions at the OSTP — which has coordinated work on the last three government climate assessments — or high-level science posts at federal agencies that work on climate change.

At the EPA, rank-and-file staff say that they haven't been told who will sign off on the science report, or how the OSTP will manage the final review process. Agency scientists told *Nature* that climate change has become taboo in their discussions with EPA leadership. The fact that agency leaders have consulted with climate sceptics has only added to the confusion.

One EPA official, who asked for anonymity because of career concerns, provided *Nature* with two lists circulating among Pruitt's team that seem to have been compiled by the Heartland Institute. One list, labelled "climate scientists", contains the names of more than 140 people, including many climate sceptics; the second names several dozen climate economists.

"It would look really bad for the administration to fight this."

The Heartland Institute would not comment on the documents, but a spokesman confirmed that Heartland has provided the EPA with names of people for a climate science 'red team'. Many agency researchers assume that Pruitt will use the lists to assemble that team, but some fear that it could be used to identify candidates for empty slots on the EPA's Board of Scientific Counselors, which advises the agency's research arm. An EPA spokeswoman declined to comment on the lists or the science report.

For the anonymous official, the question now is whether the adversarial approach embodied by the 'red team' idea will drive the Trump administration to delay the science report. "They are aware of the report," the official says. "We don't know what they are going to do." Then there is the broader national climate assessment, which will delve into questions that have profound implications for government policy, such as how coastal communities should respond to rising seas. That document is expected to go out to federal agencies this month.

Pruitt will have to be careful how he handles both documents, says Kyla Bennett, a former EPA ecologist who now works for the watchdog group Public Employees for Environmental Responsibility in North Easton, Massachusetts. The EPA could ignore the climate report's findings while implementing policies that affect the oil, gas and coal industries, which Trump has vowed to protect and promote. But if the administration pushes regulations that ignore mainstream climate science, Bennett says, it is likely to face lawsuits from environmental and science groups.

"The EPA is supposed to be using the best science out there," she says. "They can't just suddenly say the Earth is flat, CO_2 is not a pollutant and coal is the best thing for the world." ■

REPRODUCIBILITY

P-value shake-up proposed

Big names in statistics recommend tightening threshold for significance in biomedical science.

BY DALMEET SINGH CHAWLA

Science is in the throes of a reproducibility crisis, and researchers, funders and publishers are increasingly worried that the scholarly literature is littered with unreliable results. Now, a group of 72 prominent researchers is targeting what they say is one cause of the problem: weak statistical standards of evidence for claiming new discoveries.

In many disciplines, the significance of findings is judged by *P* values. They are used to test (and dismiss) a 'null hypothesis', which generally posits that the effect being tested for doesn't exist. The smaller the *P* value that is found for a set of results, the less likely it is that the results are purely due to chance. Results are deemed 'statistically significant' when this value is below 0.05.

But many scientists worry that this threshold has caused too many false positives to appear in the literature, a problem exacerbated by a practice called *P* hacking, in which researchers gather data without first creating a hypothesis to test, and then look for patterns in the results that can be reported as statistically significant.

So, in a provocative manuscript posted on the PsyArXiv preprint server on 22 July, researchers argue that *P*-value thresholds should be lowered to 0.005 for the social and biomedical sciences (D. Benjamin *et al.* Preprint at PsyArXiv http://osf.io/preprints/psyarxiv/mky9j; 2017). The final paper is set to be published in *Nature Human Behaviour*.

"Researchers just don't realize how weak the evidence is when the *P* value is 0.05," says Daniel Benjamin, one of the paper's co-lead authors and an economist at the University of Southern California in Los Angeles. He thinks

MORE ONLINE

IMAGES OF THE MONTH

NASA/JPL-CALTECH/SWRI/MSSS/JASON MAJOR

July's sharpest science shots — as selected by *Nature*'s photo team go.nature.com/2vdltOg

TOP NEWS

- Why astronomers reluctantly announced a possible exomoon discovery go.nature.com/2tueikq
- Landslide triggered rare Greenland mega-tsunami go.nature.com/2hjtsce
- Clues emerge in mystery of flickering quasars go.nature.com/2tutftt

NATURE PODCAST

The first flower; gene editing human embryos; and the quest for antimatter nature.com/nature/podcast

that claims with *P* values between 0.05 and 0.005 should be treated merely as "suggestive evidence" instead of established knowledge.

Other co-authors include two heavyweights in reproducibility: John Ioannidis, who studies scientific robustness at Stanford University in California, and Brian Nosek, executive director of the Center for Open Science in Charlottesville, Virginia.

One problem with reducing *P*-value thresholds is that it may increase the odds of a false negative — stating that effects do not exist when in fact they do — says Casper Albers, a researcher in psychometrics and statistics at the University of Groningen in the Netherlands. To counter that, Benjamin and his colleagues suggest that researchers increase sample sizes by 70%; they say this would avoid increasing rates of false negatives, while still dramatically reducing rates of false positives. But Albers thinks that, in practice, only well-funded scientists would have the means to do this.

Shlomo Argamon, a computer scientist at the Illinois Institute of Technology in Chicago, says there is no simple answer to the problem, because "no matter what confidence level you choose, if there are enough different ways to design your experiment, it becomes highly likely that at least one of them will give a statistically significant result just by chance". More-radical changes, such as new methodological standards and research incentives, are needed, he says.

Lowering *P*-value thresholds may also exacerbate the 'file-drawer problem', in which studies with negative results are left unpublished, says Tom Johnstone, a cognitive neuroscientist at the University of Reading, UK. But Benjamin says that all research should be published, regardless of *P* value.

Other scientific fields have already cracked down on *P* values — and in 2015, the journal *Basic and Applied Social Psychology* banned them. Particle physicists, who collect reams of data from accelerator experiments, have long demanded a *P*-value threshold below 0.0000003 (or 3×10^{-7}) because of concerns that a lower threshold could lead to mistaken claims, notes Valen Johnson, a statistician at Texas A&M University in College Station and a co-lead author of the paper. More than a decade ago, geneticists took similar steps to establish a threshold of 5×10^{-8} for genome-wide association studies, which look for differences between people with a disease and those without across hundreds of thousands of DNA-letter variants.

Yet other scientists have abandoned *P* values in favour of more-sophisticated statistical tools, such as Bayesian tests, which require researchers to define and test two alternative hypotheses. But not all researchers will have the technical expertise to carry out Bayesian tests, says Johnson, who thinks that *P* values can still be useful for gauging whether a hypothesis is supported by evidence. "*P* value by itself is not necessarily evil." ■

CHRISTOPHE LEPETIT/HEMIS/ALAMY

A tiny island in the Tetiaroa atoll near Tahiti is now nearly mosquito-free.

PEST CONTROL

Mosquitoes meet their match in Tahiti

Bacteria-laden insects deployed on South Pacific islands in effort to rid the region of the pests.

BY EMMA MARRIS, PAPEETE, TAHITI

The South Pacific islands have long drawn sailors and tourists seeking paradise on Earth, but biologists are now trying to make the region even more alluring. A biomedical lab in Tahiti has succeeded in nearly eradicating mosquitoes from a tiny nearby island, and researchers are gearing up to eliminate the pests from a larger island that is permanently inhabited by people.

The eventual goal is to cut off transmission routes for mosquito-borne diseases such as dengue, chikungunya and Zika, which plague the Pacific. Researchers also hope that reducing the mosquito burden will help populations of local birds. On other islands, such as Hawaii, avian malaria spread by mosquitoes can devastate bird populations.

The mosquito problem could be solved in the Society Islands — a part of French Polynesia that includes Tahiti, Moorea, Bora Bora, Huahine and Raiatea — within ten years, says Hervé Bossin, an entomologist at the mosquito lab of the Louis Malardé Institute in Paea, Tahiti, and the project's lead scientist.

He and his team plan to do this using a technique that infects mosquitoes with a specific strain of a bacterium called *Wolbachia*. About 65% of insects around the world carry *Wolbachia*, but the strains vary. If mosquitoes with different strains mate, the resulting eggs develop incorrectly and don't hatch. If there are enough of these doomed pairings, an area's mosquito population usually dies out.

But first, scientists must sort the males from the females. In a small, tidy lab on Tahiti's east coast, nestled among coconut palms and fragrant white tiare blossoms, senior technician Michel Cheong Sang pours water between two glass plates set at an angle, washing several dozen larvae of *Aedes polynesiensis* mosquitoes down between them. The larger females get stuck about halfway down. The smaller males descend a bit farther, forming a dark, wriggling band behind the glass. The low-tech method sorts more than 99% of the larvae correctly, says Bossin.

All the larvae are infected with a particular strain of *Wolbachia* — taken from a related mosquito species, *Aedes reversi* — that is not naturally present in French Polynesia. Only the males will be released in target areas to mate with wild female mosquitoes. Researchers are working at a total of five sites, most located at luxury hotel properties around ▶