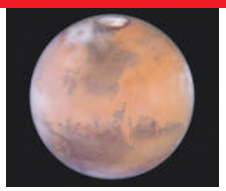


MARS IN FOCUS
For updates on Mars missions, visit our website.
www.nature.com/news/infocus/mars.html

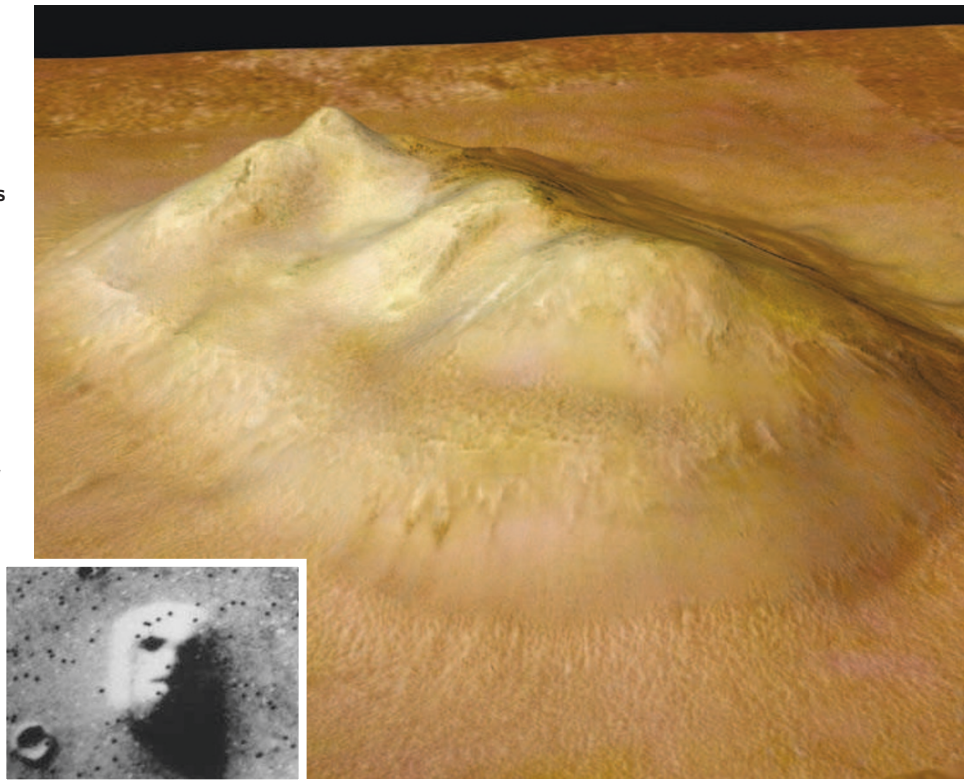
SNAPSHOT
Face facts

In a picture of Mars taken by the Viking 1 orbiter in 1976 (inset), sunlight caught this hill in a way that it made it look like a shadowy face. Here, the European craft Mars Express reveals it is nothing more than a lumpy rock — as scientists argued all along.

In 1998 and 2001, NASA's Mars Global Surveyor took high-resolution pictures of the outcrop, but failed to quell speculation by some that the feature was a relic of a martian civilization.

Gerhard Neukum, principal investigator for the Mars Express camera, says he received hundreds of e-mails asking for new pictures. "Some of them even said 'NASA is lying, we want the truth'." He says the conspiracy theorists went quiet after the release of the image shown here, constructed using data gathered on 29 December 2004 and 22 July 2006.

The hill, imaged with a resolution of 13.7 metres per pixel, is a few kilometres wide and was probably shaped by erosion. ■
Jenny Hogan



ESA/DLR/FU BERLIN/MSSS

NASA/JPL

Statistical flaw trips up study of bad stats

When two Spanish researchers checked the statistics in scientific papers from the *BMJ* and *Nature* in 2004, their results prompted a flurry of headlines and soul-searching for editors.

"Sloppy stats shame science" ran the headline in *The Economist*. "Statistical flaws revealed in top journals' papers" declared *New Scientist*. The revelation that more than a third of all *Nature* papers in 2001 contained statistical errors prompted the journal to introduce new checks on quality. At the *BMJ* (formerly the *British Medical Journal*), where one in four papers was found to contain flaws, editors ran an editorial discussing potential solutions.

Now it seems that another manuscript can be added to the pile of flawed publications: the paper by the Spanish researchers. According to an analysis published earlier this month, one of its two conclusions is invalid because it is based on inappropriate statistics. "Statistical tests should be used correctly," notes the dry conclusion of the new paper.

The test in question was employed by ecologists Emili García-Berthou and Carles Alcaraz from the University of Girona to assess the use of rounding. If rounded correctly, the final digit of a number quoted to three or four significant figures should have an equal chance of being any digit between zero and nine. García-Berthou and Alcaraz found that fours and nines appeared less frequently than expected, perhaps because researchers have an unconscious preference for rounding to five and zero, respectively (E. García-Berthou and C. Alcaraz *BMC Med. Res. Methodol.* **4,** 13; 2004).

But Monwhea Jeng, a physicist at Syracuse University in New York, points out that García-Berthou and Alcaraz employed the Kolmogorov–Smirnov test, which is used to evaluate whether the probability distributions underlying two sets of numbers differ. Jeng says this was a mistake, because the test assumes that the distributions involved are continuous; in the case of the final digit of a number, the distribution is discrete. He says a more appropriate analysis using a chi-squared test shows no evidence of rounding errors (M. Jeng *BMC Med. Res. Methodol.* **6,** 45; 2006).

"I view most of the literature as done wrong."

Two independent statisticians contacted by *Nature* say Jeng's analysis seems to be correct, and a referee report, posted on the journal's website, indicates that use of a different statistical package confirmed Jeng's conclusion. But García-Berthou insists his use of the test was appropriate, and says he will explain why in a letter to the journal. The argument could be viewed as no more than amusingly ironic, especially because Jeng's analysis does not affect García-Berthou and Alcaraz's second and perhaps more important discovery: that many *P* values in *Nature* and *BMJ* papers were wrongly calculated. But the incident does illustrate the difficulty of assessing whether statistical tests have been properly used, especially given that scientists well trained in statistics do not always agree.

"If a journal doesn't have enough expertise it's a real problem," says Steve Goodman, a medical statistician and editor at the *Annals of Internal Medicine*. He says he rarely sees papers that are free of flaws. "I view most of the literature as done wrong."

A study published this July, for example, examines the use of *P* values in paper abstracts (P. C. Gøtzsche *BMJ* **333,** 231–234; 2006). The paper concludes: "Significant results in abstracts are common but should generally be disbelieved." ■
Jim Giles

©2006 Nature Publishing Group