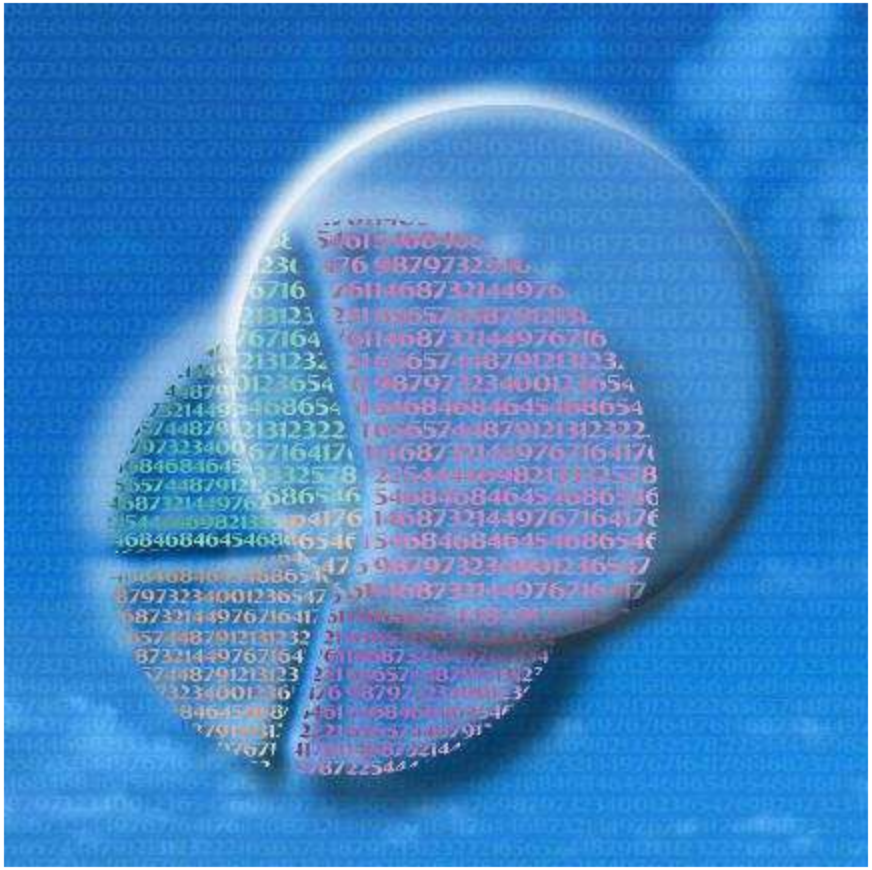


Course at the Joint Research Centre of Ispra:

‘Sensitivity analysis, sensitivity auditing and beyond’
Part on the p-test

Andrea Saltelli
Centre for the Study of the Sciences and the Humanities (SVT), University of Bergen (UIB) and
Institut de Ciència i Tecnologia Ambientals (ICTA) -Universitat Autonoma de Barcelona (UAB)

Ispra March 29-31

Where to find this talk: www.andreasaltelli.eu

= more material on my web site

= discussion time

ROYAL SOCIETY OPEN SCIENCE

rsos.royalsocietypublishing.org

Review

CrossMark click for updates

Cite this article: Colquhoun D. 2014 An investigation of the false discovery rate and the misinterpretation of *p*-values. *R. Soc. open sci.* **1**: 140216.
http://dx.doi.org/10.1098/rsos.140216

An investigation of the false discovery rate and the misinterpretation of *p*-values

David Colquhoun

Department of Neuroscience, Physiology and Pharmacology, University College London, Gower Street, London WC1 6BT, UK

"If you are foolish enough to define 'statistically significant' as anything less than *p*=0.05 then… you have a 29% chance (at least) of making a fool of yourself.

Who would take a risk like that? Judging by the medical literature, most people would. No wonder there is a problem"

Colquhoun D. 2014 An investigation of the false discovery rate and the misinterpretation of p-values. R. Soc. Open sci. 1: 140216. http://dx.doi.org/10.1098/rsos.140216

P values by way of an example

- Two groups, one with a placebo, one with the treatment
- Random allocation to groups (+more!)
- The difference *d* between the means of the two groups is tested (is it different from zero?)
- $p=0.05$ implies that if there were no effect the probability of observing a value equal to *d* or higher would be *5%*

"At first sight, it might be thought that this procedure would guarantee that you would make a fool of yourself only once in every 20 times that you do a test"

Colquhoun D. 2014 An investigation of the false discovery rate and the misinterpretation of p-values. R. Soc. Open sci. 1: 140216. http://dx.doi.org/10.1098/rsos.140216

"The classical p-value does exactly what it says. But it is a statement about what would happen if there were no true effect. That cannot tell you about your long-term probability of making a fool of yourself, simply because sometimes there really is an effect. In order to do the calculation, **we need to know a few more things**"

Colquhoun D. 2014 An investigation of the false discovery rate and the misinterpretation of p-values. R. Soc. Open sci. 1: 140216. http://dx.doi.org/10.1098/rsos.140216

A classic exercise in screening

You test positive for AIDS (one test only). Time for despair?

Only one 1 in 100,000 has AIDS in your population

The test has a 5% false positive rate

Already one can say: in a population of say 100,000 one will have AIDS and 5,000 (5% of 100,000) will test positive

➔ Don't despair (yet)

Another exercise in screening (Colquhoun 2014)

You test positive for mild cognitive impairment (MCI) (one test only). Time to retire?

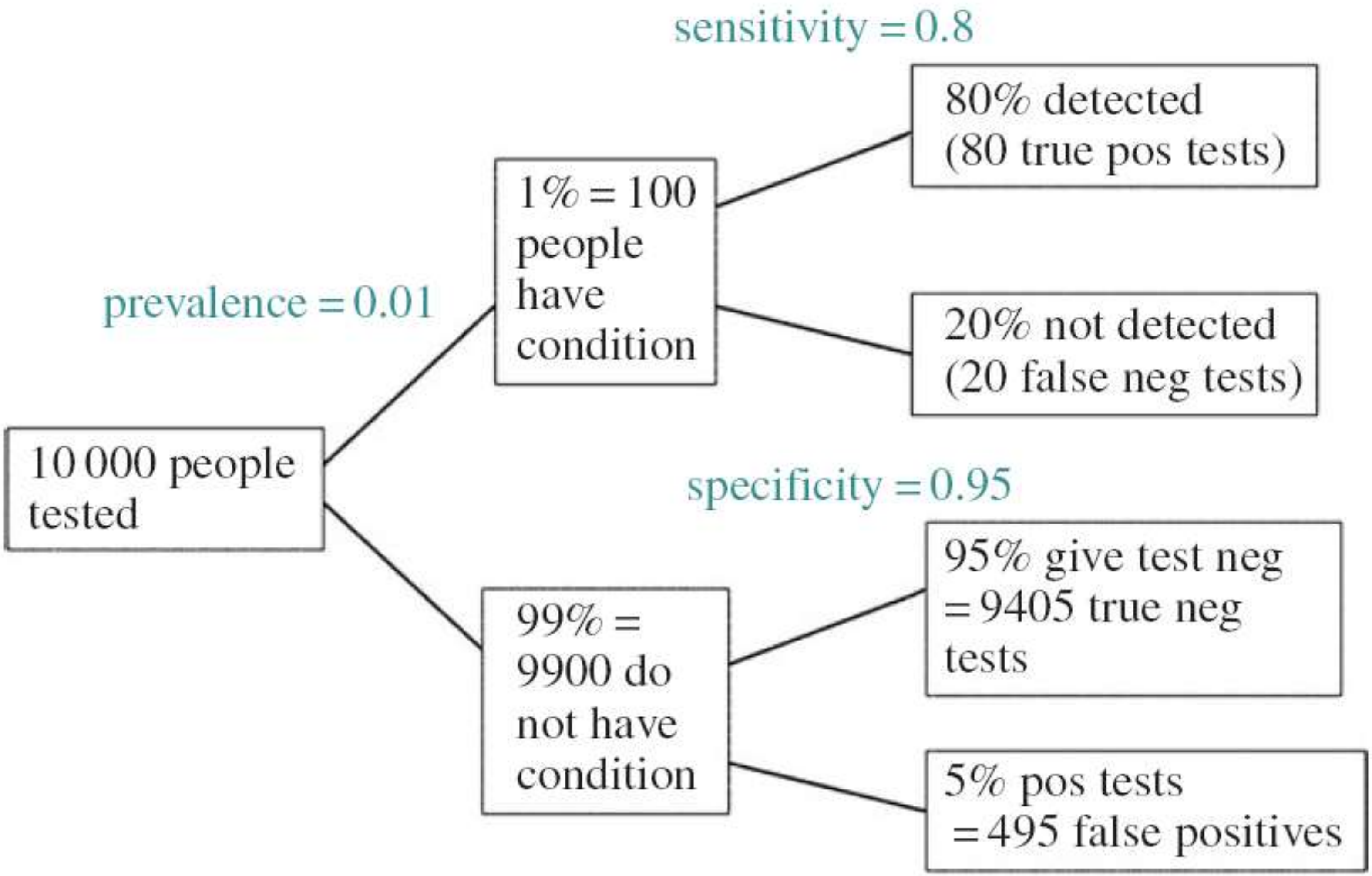
MCI prevalence in the population 1%, i.e. in a sample of 10,000 then 100 have MCI and 9,900 don't

The test has a 5% false positive rate; of the 9,900 who don't have MCI 495 test (false) positive and the remaining 9,405 (true) negative

The test does not pick all the 100 MCI but only 80; there will be 20 false negative. So we see 80+495=575 positive of which only 80 (a 14%) are true and the remaining 86% false

➔ It does not make sense to screen the population for MCI!

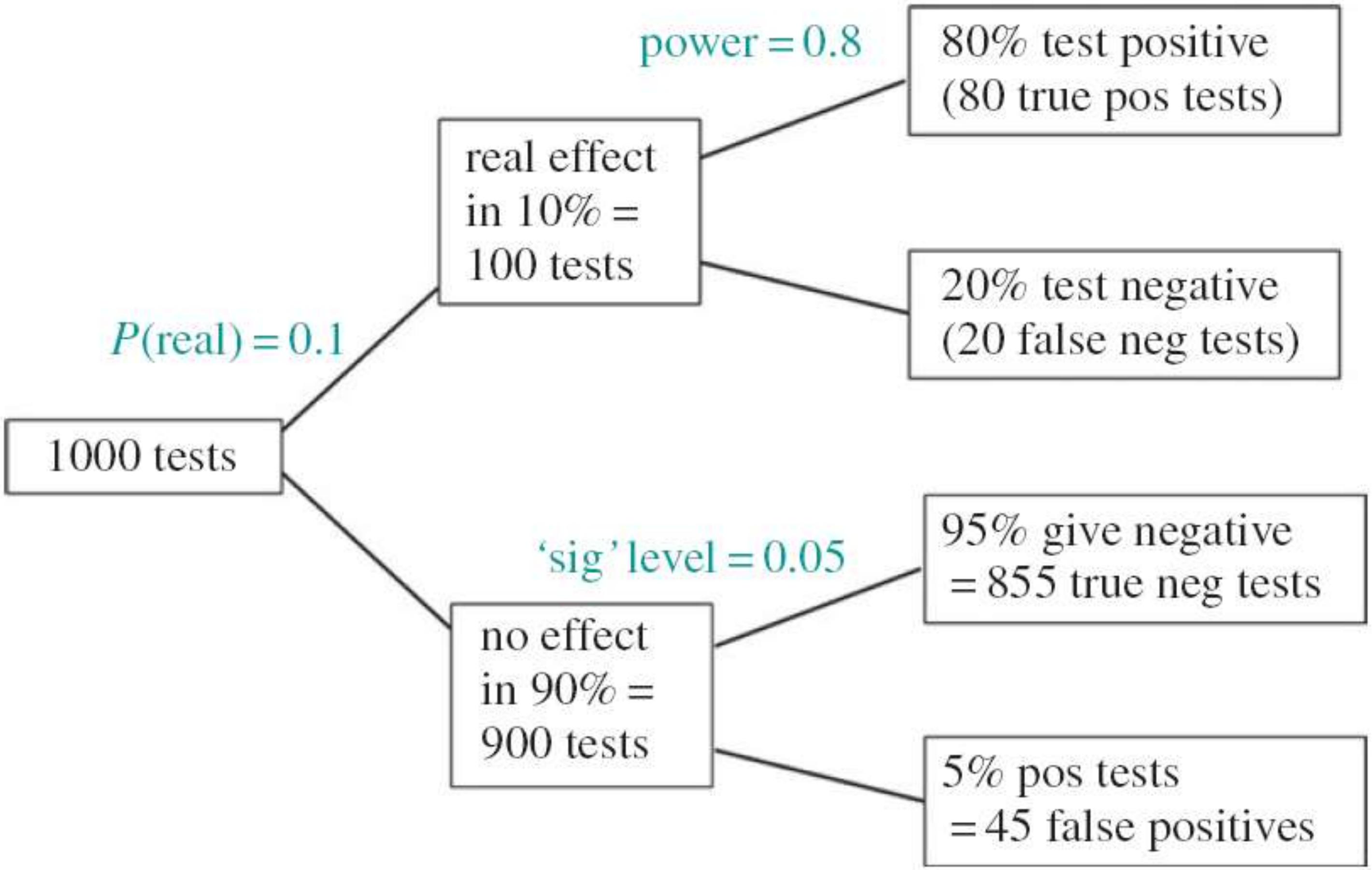
The number 86% = 495/(495+80) is our false discovery rate



Colquhoun D. 2014 An investigation of the false discovery rate and the misinterpretation of p-values. R. Soc. Open sci. 1: 140216. http://dx.doi.org/10.1098/rsos.140216

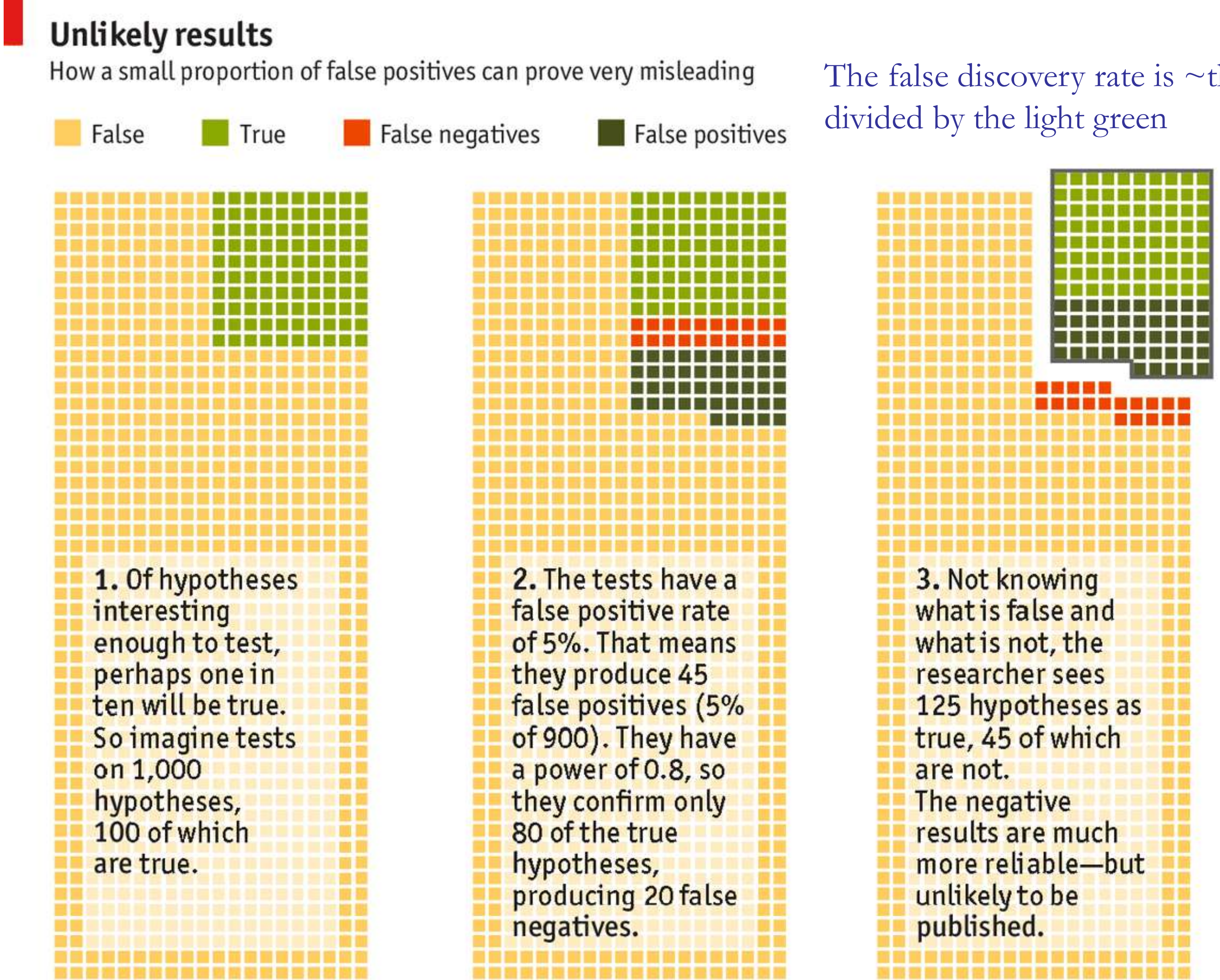
The same concept of false discovery rate applies to the problem of significance test

We now consider tests instead of individuals



Colquhoun D. 2014 An investigation of the false discovery rate and the misinterpretation of p-values. R. Soc. Open sci. 1: 140216. http://dx.doi.org/10.1098/rsos.140216

The false discovery rate is ~the dark divided by the light green



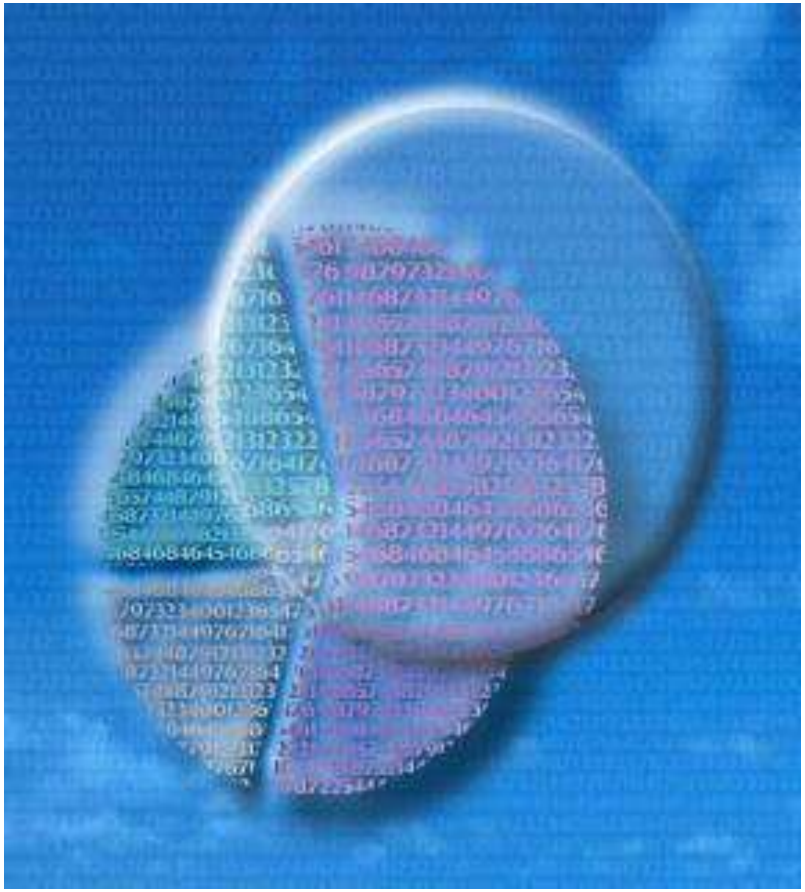
Source: *The Economist*

➔ We see 125 hypotheses as true 45 of which are not; the false discovery rate is $45/125 = 36\%$

Significance p=0.05 ➔ false discovery rate of 36%

We now know that p=0.05 did not correspond to a chance in twenty of being wrong but to one in three

How many numbers did we need to know to reach this conclusion?



END

Twitter:
@andreasaltelli