

NEWS

Researchers correct statistical flaw in high-profile paper

BY JESSICA WRIGHT

3 MAY 2017



It's every scientist's worst nightmare: A simple mistake in a high-profile study results in a large statistical error that then becomes embarrassingly public.

That happened to **Evan Eichler** and his team with a paper they published in *Nature Genetics* in February, **tying 38 new genes to autism or developmental delay**.

On 12 March, a group of 14 independent geneticists pointed out an error in a **critique they posted** on the preprint server bioRxiv. The geneticists also sent their concerns to the editors at *Nature Genetics*.

Eichler is working with the journal to correct the findings — and says he and his team are devastated.

“When we heard about this, we dropped everything,” says Eichler, professor of genome sciences at the University of Washington in Seattle. “I wish it had never happened; I wish we had caught that error; we appreciate the fact that they called it out.”

Eichler’s team sequenced 208 genes implicated in autism or developmental delay in nearly 12,000 people with either condition and combined their data with previous findings in these genes¹. This approach lifted 38 new genes above a statistical bar they had not previously reached, leading the researchers to deem them new candidates for the conditions.

But the researchers had used a flawed statistical analysis: After correcting the error, only one of the new genes retains its statistical ties to autism, according to the critique.

The public nature of the critique came as a surprise to Eichler and his team, but the authors say their intent was only to correct the record quickly, in order to prevent other scientists from pursuing the incorrect results.

“We need a faster approach to flagging for the community when something that has been published is actually significantly flawed,” says lead author **Mark Daly**, co-director of the medical and population genetics program at the Broad Institute in Massachusetts.

Daly and his colleagues had once before attempted to correct a statistical error, in a 2012 study that proposed using genetic variants to ‘diagnose’ autism. The correction was urgent because biotechnology companies **had begun using the flawed data** to create diagnostic tests. Still, it took more than a year for the letter and the correction to appear in a journal, Daly says.

His team made the commitment last year to **publish all their findings on bioRxiv** at the same time they submit to a journal. They turned to the preprint server to publish this critique as well.

The authors are also sympathetic to the study team’s plight.

“It was a mistake, and we all live in fear of making mistakes — myself in particular,” says **Bernie Devlin**, professor of psychiatry at the University of Pittsburgh, who co-wrote the critique.

Fixing flaws:

The error in the study originates in a statistical analysis intended to determine whether mutations in certain genes are more prevalent in people with autism than in the general population.

Eichler’s team looked at 208 genes that previous studies had linked to autism or developmental

delay. To boost the genes' association with autism, the team combined new mutations they found with results from previous studies.

This is where they went wrong: They analyzed their combined data as though they were only looking at 208 genes, instead of all 20,000 genes in the genome, which the previous studies had sequenced.

The error falsely boosted 28 genes, even though Eichler's team had not found any new mutations in these genes.

"We should have seen that when we went across the tables, and we just didn't — that should have been a giveaway right there," says Eichler. "It's a tough way to learn, but boy, do you learn."

The critique also noted that the researchers compared people who have autism with a database of controls from different ethnic backgrounds, and whose genomes were sequenced using different techniques.

Eichler says he is waiting for direction from the editors at *Nature Genetics* on this point.

It is unclear when the correction will be published. Eichler says he regrets the paper's title, which includes 91, the total number of autism genes the study reported. In fact, he says, sequencing data that has become available since the study's publication ups the number to 151. "By the time this all gets said and done, the list of genes will be so obsolete, people will be moving on to a gene list that's three times the size," he says.

The critique also does not take away from the study's other findings.

For example, the researchers looked for genes mutated more often in people with autism than in those with intellectual disability, and vice versa — an approach others say is valuable.

"There are elements of this paper that were very good and that also shouldn't be lost," says Daly.

Eichler is considering updating the findings on bioRxiv before the official correction appears in the journal, as the latter may take some time.

REFERENCES:

1. Stessman H.A. *et al. Nat. Genet.* **49**, 515-526 (2017) [PubMed](#)