

NEWS

Information loss may weaken autism genetic scores

BY LAURA DATTARO

13 MAY 2022

Genetic predictions of an individual's chances of having autism may often **miss the contributions from top variants** associated with the condition, according to findings presented yesterday at the **2022 International Society for Autism Research** annual meeting. (Links to abstracts may work only for registered conference attendees.)

With hundreds of genes and thousands of variants linked to autism, some researchers are working to create reliable '**polygenic scores**.' Among them are **Kelly Benke**, assistant scientist at Johns Hopkins University in Baltimore, Maryland, who enlisted Michael Yao, a high school student volunteer, to check her analysis. Yao, who presented the work, discovered that despite using the best protocols, they had inadvertently dropped some variants strongly linked to autism.

"There is information loss in these scores," Benke says.

To test how widespread the problem might be, Benke and her team searched three datasets — the **Early Autism Risk Longitudinal Investigation** (EARLI) study, the **Infant Brain Imaging Study** (IBIS) and the **Markers of Autism Risk in Babies - Learning Early Signs** (MARBLES) study — for 88 variants, called single nucleotide polymorphisms (SNPs), previously identified in a large genome-wide association study. Even after standard data correction, one analysis missed six of the variants.

"The genetics community doesn't report this," Benke says. "You just say, 'This is how many SNPs went into the score.' That tells you nothing about whether you've lost important information and how much."

Before the researchers cleaned the data — a common process called imputation — just 28 of the 88 variants appeared in MARBLES, 52 in IBIS and 54 in EARLI, they found. After imputation, however, they could identify all 88 variants in IBIS and EARLI, and 82 in MARBLES.

The MARBLES study uses a microarray chip that can search 1 million genetic sites for SNPs, whereas EARLI and IBIS use a newer one — the largest available — that can search 5 million sites. But the 5M chip costs about \$500 per genome, Benke says, making it prohibitively expensive for many studies.

The team also created a new metric to score how well a given polygenic score incorporates variants most strongly linked to autism. Out of a top score of 1, IBIS scored 0.9, EARLI 0.88 and MARBLES 0.87.

The metric could be useful in future publications of polygenic scores, Benke says. She and her colleagues are considering creating a free tool to help researchers make the calculation, but in the meantime the necessary information is freely available, she says.

*Read more reports from the **2022 International Society for Autism Research** annual meeting.*

Cite this article: <https://doi.org/10.53053/OWBS1849>