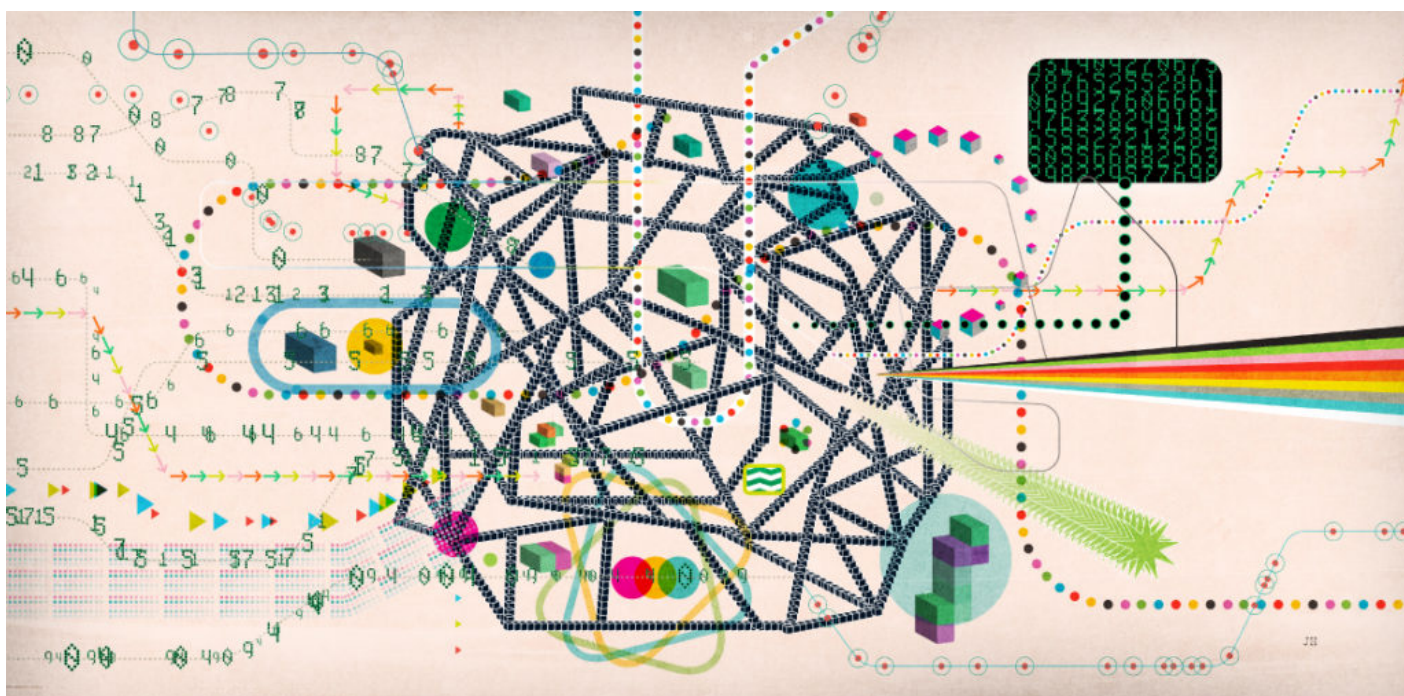


NEWS

Algorithm automates efforts to estimate autism prevalence

BY RACHEL ZAMZOW

30 JANUARY 2017



A computer program can accurately estimate the **prevalence** of autism from children's medical and school records, suggests a new report¹. The tool could help automate efforts to estimate autism rates over time, cutting down the enormous time and costs now required.

In 2000, the U.S. Centers of Disease Control and Prevention (CDC) set up the **Autism and Developmental Disabilities Monitoring (ADDM) Network** to track autism prevalence. The network consists of sites in 11 states that collect medical and school records for a subset of 8-year-old children in part of each state. Every two years, clinicians comb through the records, looking for descriptions that suggest a child has autism.

This manual method has become more labor-intensive with the rise in autism diagnoses. Clinicians at one ADDM Network site went from analyzing records for 1,152 children in 2000 to sifting data for 9,811 children in 2010.

“The amount of information that is available that goes through the surveillance system has really just exploded over this time,” says lead researcher **Matthew Maenner**, an epidemiologist with the CDC.

Clinicians spend about an hour reviewing records for each child. A second clinician may be called in to review ambiguous records. The new algorithm, described 21 December in *PLOS One*, can process data for more than 1,000 children in a few minutes.

Word search:

Maenner and his colleagues developed the algorithm by first teaching it to identify words and phrases important for determining whether a child has autism. The algorithm selected 90 terms, such as 'autism,' 'social interaction' and 'eye contact,' that significantly influence how well the program agrees with clinicians.

The researchers trained the algorithm with data from the Georgia site of the 2008 ADDM Network. They taught the algorithm to spot differences in the distribution of the key terms in the records of children with autism and those of children without it.

The team then tested the algorithm, once again using data from the Georgia site, but this time from 2010. The algorithm classified 708 of 1,450 children as having autism, putting the prevalence at 1.46 percent — close to the clinicians' estimate of 1.55 percent.

The software matched the clinicians' classifications 86.5 percent of the time. Clinicians in the 2010 network agreed with one another's assessments 90.7 percent of the time.

This difference is too small to be of concern, says **Eric Fombonne**, professor of psychiatry at Oregon Health and Science University, who was not involved in the study. But the algorithm may not capture some subgroups of individuals, such as girls, as well as clinicians do, he says.

Streamlining surveillance:

The children the algorithm missed had fewer evaluations or were evaluated at a later age, on average, than those it detected. The program was also error-prone when classifying children that the clinicians themselves had trouble sorting.

Missing a few cases of autism would not greatly affect prevalence estimates, however, says **Maureen Durkin**, professor of population health sciences and pediatrics at the University of

Wisconsin-Madison, who was not involved in the study. “For the purpose of surveillance, they could probably begin to use this approach fairly soon,” she says.

Still, the researchers should first test the method using data from other ADDM sites, says Fombonne. Autism rates differ between states, and it’s not yet clear whether the algorithm would work as well with data from other sites. “That could really affect the performance of their method,” he says.

Maenner stresses that the algorithm is not intended to replace clinicians. “We’re not anywhere near flipping a switch and saying, ‘Okay, computers are taking over now,’” he says.

But automation could help to streamline autism surveillance by serving as an initial filter of children’s records, he says. The algorithm could make classifications when the data are clear and pinpoint records that need further review.

REFERENCES:

1. Maenner M.J. *et al.* *PLOS One* **11**, e0168224 (2016) [PubMed](#)