IMPORTANCE The origins and development of autism spectrum disorder (ASD) remain unresolved. No individual-level study has provided estimates of additive genetic, maternal, and environmental effects in ASD across several countries.

OBJECTIVE To estimate the additive genetic, maternal, and environmental effects in ASD.

DESIGN, SETTING, AND PARTICIPANTS Population-based, multinational cohort study including full birth cohorts of children from Denmark, Finland, Sweden, Israel, and Western Australia born between January 1, 1998, and December 31, 2011, and followed up to age 16 years. Data were analyzed from September 23, 2016 through February 4, 2018.

MAIN OUTCOMES AND MEASURES Across 5 countries, models were fitted to estimate variance components describing the total variance in risk for ASD occurrence owing to additive genetics, maternal, and shared and nonshared environmental effects.

RESULTS The analytic sample included 2,001,631 individuals, of whom 1,027,546 (51.3%) were male. Among the entire sample, 22,156 were diagnosed with ASD. The median (95% CI) ASD heritability was 80.8% (73.2%-85.5%) for country-specific point estimates, ranging from 50.9% (25.1%-75.6%) (Finland) to 86.8% (69.8%-100.0%) (Israel). For the Nordic countries combined, heritability estimates ranged from 81.2% (73.9%-85.3%) to 82.7% (79.1%-86.0%). Maternal effect was estimated to range from 0.4% to 1.6%. Estimates of genetic, maternal, and environmental effects for autistic disorder were similar with ASD.

CONCLUSIONS AND RELEVANCE Based on population data from 5 countries, the heritability of ASD was estimated to be approximately 80%, indicating that the variation in ASD occurrence in the population is mostly owing to inherited genetic influences, with no support for contribution from maternal effects. The results suggest possible modest differences in the sources of ASD risk between countries.
Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by impairments in social interaction and communication and the presence of restricted interests and repetitive behaviors.\(^1\)\(^2\)

Autism spectrum disorder has both genetic and environmental origins. Research into the genetic origins of ASD has consistently implicated common and rare inherited variation (heritability). However, evidence shows that there are other, noninherited, genetic influences that could be associated with variation in a trait.\(^3\) Given the prenatal origins of ASD, an important source of such genetic influences could be maternal effects.\(^4\) The term *maternal effects* is used to describe the association of a maternal phenotype with ASD in offspring (ie, the noninherited genetic influences originating from mothers beyond what is inherited by the offspring). Maternal effects have been associated with a substantial proportion of the variation in several traits associated with ASD, including preterm birth\(^5\) and intelligence quotient.\(^6\) Research on nongenetic origins has frequently pointed to a role for environmental exposures unique to different family members (nonshared environment), an example of which is cesarean delivery.\(^7\) In contrast, contribution from environmental exposures that make family members similar (ie, shared environment), has been uncertain.\(^8\)

A meta-analysis of twin studies estimated heritability to be in the range of 64% to 91%,\(^8\)\(^9\) and 3 population-based studies from Sweden recently estimated the heritability of ASD to be 83%,\(^8\) 80%,\(^4\) and 66%.\(^10\) Among those earlier heritability calculations from twin and family studies (eTable 1 in the Supplement), a single study has estimated maternal effects,\(^4\) reporting modest, if any, contribution to ASD. Estimates of the contribution of shared environment range from 7% to 35%,\(^8\) but multiple studies estimate the contribution to be zero.\(^4,9,11,12\) Thus, although the origin and development of ASD has been investigated for half a century, it remains controversial.

The current study was designed to rigorously determine the contribution of various genetic and nongenetic origins hypothesized for ASD. We aimed to estimate the heritability together with maternal effects and shared and nonshared environmental effects of ASD. To achieve this aim, we used what is to our knowledge the largest-ever dataset for population-based epidemiologic autism research to date containing family data from 5 countries and generalized linear mixed models designed to quantify the variation in ASD liability owing to genetic and environmental origins using information about ASD in family members of varying levels of relatedness (eg, siblings, cousins).\(^4\) We aimed to examine the consistency of these estimates by contrasting the results across 5 different countries. Data were analyzed for each country and in an individual-level pooled analysis for Nordic countries to maximize precision.

### Methods

This study was approved by the Danish Data Protection Agency, the Danish National Board of Health, the institutional review boards of the University of Haifa and the Helsinki Ethics Committee, the ethics committee of the Finnish National Institute for Health and Welfare and hospital district of Southwest Finland, the Swedish Ethical Review Board Stockholm, the Department of Health Western Australia Human Research Ethics Committee, and the Institutional Review Board of the New York State Psychiatric Institute. Those bodies waived the need for informed consent because the study data were fully deidentified.

### Study Population

The study population comprised all singleton live births in Denmark, Finland, Sweden, Israel, and Western Australia. For Denmark, Sweden, Finland, and Western Australia, we included all births between January 1, 1998, and December 31, 2011, in Western Australia. Data on 3-generational family linkages, allowing for the identification of parents, siblings, and cousins, were available for all sites (eTable 2 in the Supplement). Data were analyzed from September 23, 2016, to February 4, 2018.

### Outcome and Covariates Information

Outcome and covariate information (birth year, sex) were provided by government-maintained national health registries.\(^14\) Different diagnostic systems were used across sites (eTable 2 in the Supplement) and the diagnostic codes for ASD were harmonized accordingly (eTable 3 in the Supplement). Data from Israel were only used for ASD analyses because autistic disorder (AD) diagnosis was not available. Case ascertainment and reliability and validity of registry-reported diagnoses have been published previously.\(^15,16\)

### Statistical Analysis

**Analytic Samples and Statistical Models**

We chose a multigenerational family design because the strength of genetically induced correlation for a given trait varies by the degree of relatedness between relatives, which...
allows additive genetic (heritability), maternal, and shared environmental effects to be estimated. A residual term, which is commonly interpreted as nonshared environmental effects, can also be estimated.\(^{17}\) Thus, nonshared environment will not be an estimate of nonshared environmental effects only but will also include contributions from sources not explicitly represented by any model parameters, including gene-environmental correlations and part(s) of any gene-environmental interactions. We chose to use the term non-shared environment for this term to not diverge from earlier publications in this field. The data preparation and modeling approach has been described in detail elsewhere\(^4\) (eAppendix 2 in the Supplement). Briefly, we used the 3-generational data sources to construct families that vary by genetic relatedness and therefore are informative for genetic modeling. These included full siblings and cousins related through their mothers (maternal parallel cousins [mPCs]), or cousins of other relationships. Detailed examples illustrating how families were created are provided in eFigure 1 through eFigure 5 in the Supplement.

Liability threshold models were fitted using the structured family data to decompose the variance in liability to ASD into additive genetic (A), maternal (M), and shared (C) and nonshared (E) environmental components. Three nested models were fitted: (I) the AE model: additive genetic (A) + nonshared environment (E); (II) the ACE model: additive genetic (A) + maternal effect (M) + nonshared environment (E); and (III) the ACME model: additive genetic (A) + maternal effect (M) + shared environment (C) + maternal effect (M) + nonshared environment (E). Figure 1 and eTable 4 in the Supplement show how the different components of variance are derived from different correlations between family members representing additive genetic, shared environment, and non-shared environment. These components can be derived from twins (monozygotic vs dizygotic); full siblings vs half siblings; and full siblings vs cousins. The M component can only be derived from full siblings vs mPCs vs cousins of other relationships and full siblings vs maternal vs paternal half siblings. When involving half siblings, correlation of shared environmental effect was set at 1 for maternal half siblings and 0 for paternal half siblings; an assumption behind this decision is that children lived with their mothers after a divorce and separation from the father, an assumption frequently made in these types of models; this is also the rationale for excluding half siblings in our analytic sample. Liability models using full siblings and cousins to estimate A, C, and E have been used earlier and for other outcomes (eg, skin cancer,\(^{18,19}\) M component on preterm birth\(^5\) and preeclampsia,\(^{20,21}\) and comorbidity of 2 diseases).\(^{22}\) To calculate CIs, instead of relying on the assumption that estimates follow an asymptotic normal distribution, we calculated 2-sided 95% CIs using profile likelihood methods.\(^{23}\)

Because Denmark, Finland, and Sweden share similar health, ascertainment, and diagnostic systems, and because of sample size considerations, the primary pooled analysis focused on the Danish, Finnish, and Swedish samples. First, we fitted country-specific models for Denmark, Finland, and Sweden. Then, a model was fitted for the individual-level combined data from Nordic countries. Next, we included 2 smaller samples from Israel and Western Australia. For each country, the categorical covariates sex (male vs female) and birth year cohort (2006-2011 vs 2000-2005 for Israel and 2003-2007 vs 1998-2002 for all other countries) were included as fixed factors.

We used the ViPAR\(^24\) application running R statistical software\(^25\) version 3.1.2 (R Foundation) on a Linux RedHat version 6.0 64-bit server (Hewlett Packard) or all calculations except for Israel, for which we used R version 3.4.0 on a Linux/ GNU 64-bit server through Ubuntu 16.04 (Canonical Ltd). Further details about the choice of statistical software and analysis packages are provided in eAppendix 2 in the Supplement. All tests of statistical hypotheses were done on a 2-sided 5% level of significance.
Sensitivity and Complementary Analyses

To delineate a more impaired subtype within ASD, we repeated all analyses for AD, a diagnostic category present in the ninth and tenth revisions of the International Classification of Diseases, Ninth Revision, Clinical Modification and the revised third and the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders. In addition, we performed an extensive set of analyses to test the robustness of our results. We performed 3 sensitivity analyses. For Finland and Western Australia, which had a small number of concordance pairs, we re-fitted the ACE model using half siblings instead of cousins. Because a lower prevalence could affect the heritability estimates, we used data simulation reducing the number of ASD cases in the Swedish analytic sample to approximate prevalence rates in Finland and refitted the ACE model. To illustrate the model robustness, we plotted the country-specific likelihood functions of additive genetic (A) and shared environmental (C) effect for the ACE models. To test the robustness of our results, because the analytic sample used for the statistical models did not include the entire study cohort, we performed additional analyses to ensure that the analytic sample was representative of the study cohort. We also compared age-specific outcome ascertainment and follow-up pattern between countries by constructing country-specific inverse Kaplan-Meier curves for ASD and AD assuming independent censoring (eAppendix 3 in the Supplement).

Results

The analytic sample included 2,001,631 children, of whom 1,027,546 (51.3%) were male, from 680,502 families followed up to age 16 years. Of these, 22,156 (1.11%) children were diagnosed with ASD (Denmark, 7,580; Finland, 2,968; Sweden, 10,563; Israel, 490; Western Australia, 555) (Table 1). Outcome ascertainment information across sites can be found in eTable 2 and eTable 3 in the Supplement, which provide information about the diagnostic codes used by the different countries; eTable 9 provides information about the underlying population from which we selected the analytic sample used in the calculations.

Nordic Countries: Country-Specific and Combined Estimates

Country-specific point estimates of additive genetic effect (narrow-sense heritability) ranged between 80.7% (95% CI, 74.1%-85.4%) (Denmark) and 84.1% (95% CI, 79.7%-88.1%) (Sweden) for the AE model; 52.8% (95% CI, 29.7%-76.3%) (Finland) and 84.8% (95% CI, 76.2%-88.7%) (Sweden) for the ACE model; and 50.9% (95% CI, 25.1%-75.6%) (Finland) and 81.1% (95% CI, 69.9%-86.7%) (Sweden) for the ACME model. The heritability estimates for the Nordic pooled sample were 82.7% (95% CI, 79.1%-86.0%) for the AE,
82.2% (77.2%-85.9%) for the ACE model, and 81.2% (95% CI, 73.9%-85.3%) for the ACME model (Table 2; Figure 2).

Country-specific point estimates of maternal effects ranged between 0.4% and 1.6% in the ACME model, and the Nordic pooled sample estimate was 0.5%, but in all models the 2-sided 95% CIs included zero (Table 2; eFigure 12 in the Supplement).

Country-specific point estimates of shared environmental effects ranged between 0.0% (95% CI, 0.0%-4.6%) (Denmark) and 14.5% (95% CI, 4.5%-29.2%) (Denmark) in the ACE model and 0.0% (95% CI, 0.0%-3.7%) (Sweden) and 14.0% (95% CI, 0.0%-28.6%) (Finland) in the ACME model. The estimates from the Nordic pooled sample ranged between 0.2% (95% CI, 0.0%-2.3%) for the ACE and 0.3% (95% CI, 0.0%-2.2%) for the ACME model; CIs included zero (Table 2; eFigure 12 in the Supplement).

Country-specific point estimates for nonshared environmental effects ranged between 16.0% (95% CI, 11.9%-20.4%) (Sweden) and 19.3% (95% CI, 14.6%-25.9%) (Denmark) in the AE model; 15.1% (95% CI, 11.1%-21.1%) (Sweden) and 32.8% (95% CI, 17.6%-51.1%) (Finland) in the ACE model; and 17.5% (95% CI, 12.5%-22.5%) (Sweden); and 33.6% (95% CI, 17.6%-53.9%) (Finland) for the ACME model. The Nordic pooled sample estimates were 17.3% (95% CI, 14.1%-21.7%) for the ACE model; 17.7% (95% CI, 13.9%-26.5%) for the ACME model; and 18.1% (95% CI, 14.1%-21.7%) for the ACME model (Table 2; eFigure 13 in the Supplement).

Israel and Western Australia
Owing to sample size limitations for Israel and Western Australia, only the ACE model could be fitted for these 2 countries. Results for Israel for additive genetic effect (86.8% [95% CI, 69.8%-100.0%]) were similar to those from Denmark and Sweden, whereas results from Western Australia resembled those from Finland (additive genetic effect 53.8% [28.5%, 74.3%]) (Table 3; eFigure 6 in the Supplement).

Table 2. Autism Spectrum Disorder: Estimated Variance Components* and Associated 2-Sided 95% CI for Denmark, Finland, Sweden, and Nordic Countries Combinedb

| Model and Population | Random Effects (95% CI) |  |
|----------------------|-------------------------|--|---|---|---|
|                      | Additive Genetic (A)    | Shared Environment (C) | Maternal (M) | Nonshared Environment (E) |
| Model 1: A + E       |                         |                         |              |                            |
| Country specific     |                         |                         |              |                            |
| Denmark              | 80.7 (74.1-85.4)        | NA                      | NA           | 19.3 (14.6-25.9)           |
| Finland              | 80.8 (73.2-85.5)        | NA                      | NA           | 19.2 (14.5-26.8)           |
| Sweden               | 84.1 (79.7-88.1)        | NA                      | NA           | 16.0 (11.9-20.4)           |
| Nordic countries     | 82.7 (79.1-86.0)        | NA                      | NA           | 17.3 (14.0-20.9)           |
| Model 2: A + C + E   |                         |                         |              |                            |
| Country specific     |                         |                         |              |                            |
| Denmark              | 80.4 (71.3-86.8)        | 0.0 (0.0-4.6)           | NA           | 19.6 (12.9-25.3)           |
| Finland              | 52.8 (29.7-76.3)        | 14.5 (4.5-29.2)         | NA           | 32.8 (17.6-51.1)           |
| Sweden               | 84.8 (76.2-88.7)        | 0.1 (0.0-3.4)           | NA           | 15.1 (11.1-21.1)           |
| Nordic countries     | 82.2 (77.2-85.9)        | 0.2 (0.0-2.3)           | NA           | 17.7 (13.9-26.5)           |
| Model 3: A + C + M + E |                       |                         |              |                            |
| Country specific     |                         |                         |              |                            |
| Denmark              | 78.9 (65.4-84.3)        | 0.1 (0.0-4.8)           | 0.4 (0.0-6.9) | 20.6 (14.6-26.5)           |
| Finland              | 50.9 (25.1-75.6)        | 14.0 (0.0-28.6)         | 1.6 (0.0-13.4)| 33.6 (17.6-53.9)           |
| Sweden               | 81.1 (69.9-86.7)        | 0.0 (0.0-3.7)           | 1.4 (0.0-6.6) | 17.5 (12.5-22.5)           |
| Nordic countries     | 81.2 (73.9-85.3)        | 0.3 (0.0-2.2)           | 0.5 (0.0-3.5) | 18.1 (14.1-21.7)           |

Abbreviation: NA, not applicable.
* All estimates are recalculated to fraction of variation explained, the proportion of total variance explained by each random effect (eAppendix 2 in the Supplement).
\[ b \] Nordic countries combined are Denmark, Finland, and Sweden.
Table 3. Autism Spectrum Disorder: Estimated Variance Components* and Associated 2-Sided 95% CIs for Israel and Western Australia

<table>
<thead>
<tr>
<th>Population</th>
<th>Random Effects (95% CI), % for Model 2: A + C + E</th>
<th>Nonshared Environment (E)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Israel</td>
<td>Additive Genetic (A) 86.8 (69.8-100.0)</td>
<td>Shared Environment (C) 0.1 (0.0-5.4)</td>
</tr>
<tr>
<td>Western Australia</td>
<td>53.8 (28.5-74.3)</td>
<td>25.2 (8.8-49.2)</td>
</tr>
</tbody>
</table>

* All estimates are recalculated to fraction of variation explained, the proportion of total variance explained by each random effect (eAppendix 2 in the Supplement).

Sensitivity Analyses
Because Finland and Western Australia showed reduced heritability estimates and an increased contribution from shared environment to ASD compared with the remaining countries (Table 2, eFigure 6 in the Supplement), we executed a series of sensitivity analyses. Finland and Western Australia had a small number of concordant cousin pairs, and we therefore refitted the ACE model using half siblings instead of cousins. The estimates from the ACE model for Finland and Western Australia using this approach were more similar to those observed for the other countries. For example, for Finland, heritability was estimated to 70.6%; shared environment estimated, 9.4%; and nonshared environment, 20.0% (eTable 5 in the Supplement).

Using a simulation approach, we reduced the Swedish ASD prevalence to approximate the level in Finland (Sweden simulation: 6.91 ASD cases per 1000 vs Finland, 6.89 cases per 1000). The heritability estimate from this simulation was more similar to that of Finland (simulation: 62.0% [95% CI, 54.1%-70.9%] vs Finland above, 50.9% [95% CI, 25.1%-75.6%]), but the shared environmental effect was 0.7% (95% CI, 0.0%-6.8%) compared with Finland (14.0% [95% CI, 0.0%-28.6%]) (eTable 6 in the Supplement). Taken together, the sensitivity analyses indicate that a random underascertainment of cases may underestimate the true heritability and increase the observed shared environment component. This is also reflected in the cumulative probability for diagnosis. Denmark and Sweden both have similar increasing cumulative probability of ASD and AD up to age 16 years, whereas the ascertainment in Western Australia and Finland is predominantly at younger ages (ASD, eFigure 10 in the Supplement; AD, eFigure 11 in the Supplement).

Complementary Analyses
We estimated genetic and environmental contributions to AD for Denmark, Finland, and Sweden. The Western Australian sample was too small, and Israel reported ASD without subtypes. Estimates of genetic, maternal, and environmental effects were similar with ASD; additive genetic effect, ACE model: 84.6% (95% CI, 69.7%-88.7%) (Denmark), 72.7% (95% CI, 54.2%-81.0%) (Finland), 76.3% (95% CI, 62.3%-83.0%) (Sweden); maternal effects for the Nordic combined sample, 0.6% (95% CI, 0.0%-4.9%), and nonshared environment ACE model, 15.1% (95% CI, 10.6%-21.5%) (Denmark), 16.7% (95% CI, 13.0%-24.5%) (Finland), 21.9% (95% CI, 16.0%-30.6%) (Sweden) (eTable 7, eTable 8, eFigure 7, eFigure 14, and eFigure 15 in the Supplement).

Demographic and clinical characteristics of the analytic cohort were representative of the corresponding populations (Table 1 vs eTable 9 in the Supplement). By visual inspection, (eFigure 8 and eFigure 9 in the Supplement), there were no differences in ASD to AD rate (cases per 1000) between the cohort population and the analytic sample. Similarly, comparisons did not reveal any differences between siblings and cousins with respect to parental age, interpregnancy interval, parental education, or parental psychiatric history (eTable 10 in the Supplement). In the analytic sample, comparisons across countries showed that siblings and cousins were similar with respect to sex ratio, AD proportion (percentage of ASD), family size, and differences in age between sibling or cousin pairs (eTable 11 in the Supplement).

All estimates of variance components were recalculated to fraction of variation explained for comparison across countries, because a number of raw estimates were associated with factors such as sample size and outcome rate (eTable 12 and eTable 13 in the Supplement). The plotted likelihood functions for the ACE model show the support in the data for estimating the variance components and the 2-sided 95% CIs (eFigure 16 in the Supplement). Estimates of fixed parameters, sex, and birth cohort for country-specific analyses and country (Denmark as reference) for the combined analyses are reported for the ACE model, which is applicable for all countries (eTable 14 and eTable 15 in the Supplement).

Discussion
The present study evaluated the contribution of various genetic and nongenetic factors to ASD risk. We estimated heritability together with maternal effects and shared and nonshared environment on ASD risk using population-based datasets from 5 countries from what is to our knowledge the largest family-based database for autism research to date. The current study results provide the strongest evidence to our knowledge to date that the majority of risk for ASD is from genetic factors. Nonshared environmental factors also consistently contribute to risk. In the models that combined data from the 3 Nordic countries, the genetic factors explained at least 73.9% of the variability in risk, and nonshared environment at most 26.5% based on the lower and upper bounds of the respective 95% CIs. These results are similar to those of recent population-based cohorts4,9,10 as well as a recent metaanalysis of twin studies,8 which estimated heritability in the range of 64% to 91%.

When we estimated the maternal effect, however, its association with variation in risk for ASD (and AD) was nonexistent or minimal. This corroborates a previous analysis of the data from Sweden.4 The importance of this finding lies in the insight it provides for understanding the risk factors associated with ASD. The absence of M effects indicates that there...
is no strong evidence of a maternal effect, driven by genetic factors shared between sisters, associated with the risk of ASD. Proposed maternal risk factors for ASD such as obesity\textsuperscript{7,26,27} do not map directly to components in the current model. The mechanisms through which such risk factors operate can be better understood using other study designs and analytic approaches, including animal models and epidemiologic studies examining specific risk factors.

Like most studies,\textsuperscript{8,11} overall shared environmental factors contributed minimally to the risk of ASD. However, there was variation in the results among samples. Although the 2 largest samples (Denmark, Sweden) did not support shared environmental influences, 2 other samples (Finland, Western Australia) did. Some variation is expected owing to population-level differences.\textsuperscript{28} Shared environmental factors in this study may reflect variables or processes that make members of the same family similar beyond genetic factors. Indeed, studies have suggested differences in genetic population structure between Finland and the rest of Europe,\textsuperscript{29,30} which could partially explain the increased shared environment and decreased heritability estimates for ASD in Finland. In addition, the statistical models are sensitive to small sample sizes because they rely on differences in the rate of concordant sibling-cousin pairs to estimate the shared environment component, as demonstrated in our sensitivity analyses. Bias owing to small sample size, ascertainment bias, or both, could lead to elevated estimates of shared environmental effect.

In the bulk of the analyses, results for AD and ASD were similar: additive genetic effects accounted for the largest influence on liability, followed by contribution from nonshared environment, with little evidence for maternal or shared environment effects. The pooled estimates for ASD were the same as for AD. Results for Israel were similar to those from Denmark and Sweden, whereas results from Western Australia resembled those from Finland. We believe that these patterns of results add support to the hypothesis that severity maps onto the load of liability factors.\textsuperscript{4} Autistic disorder was part of the 9th and 10th revisions of the International Classification of Diseases and the revised 3rd and the 4th edition of the Diagnostic and Statistical Manual of Mental Disorders coding system used for diagnoses in this study but is not part of the current Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, which does not separate ASD and AD.

The major strength of this study is the use of multiple large population-based samples with individual-level data in 3-generation pedigrees. Our data were based on prospective follow-up and health systems with equal access. This approach, following all participants from birth using population registries, avoids bias owing to self-report and retrospective collection of data and reduces selection biases owing to disease status or factors such as parental education. In addition to providing exceptional statistical power, the study directly addresses the concern of lack of replication in research findings\textsuperscript{31,32} replicating results across 5 countries and health systems.

Most ASD heritability studies have used twins.\textsuperscript{6} Although twins are important for etiologic studies, also including non-twin siblings provide greater generalizability and simultaneously improve precision. A study from Sweden\textsuperscript{9,33} used twins as well as full siblings and maternal and paternal half siblings. In a later study,\textsuperscript{4} with data overlapping with the current Swedish sample, this approach was extended to include cousins. In the current study, we now also distinguish maternal parallel cousins from cousins of other relationships to estimate maternal effects. Twin studies rely on very specific assumptions about genetic and environmental correlations. Applying different study designs make results less sensitive to these crucial assumptions. There are also different statistical techniques to estimate the underlying (likelihood-based) models. Some studies\textsuperscript{8,9,33} used structural equation models, others used likelihood estimation\textsuperscript{14,20} or calculated tetrachoric correlations.\textsuperscript{9,10} These methods and approaches come with their own strengths and limitations, but taken together they provide a more robust description of the underlying factors. Finally, there are also approaches using genetic markers (single-nucleotide polymorphisms) to estimate heritability.\textsuperscript{10,34} Although these make a valuable contribution, currently they provide only a lower bound for heritability.\textsuperscript{35} Furthermore, multiple studies have used questionnaire-based symptom checklists. Using contemporary clinical diagnoses adds to generalizability and helps to avoid biases that could be embedded in subclinical diagnoses. Finally, most previous heritability studies come from a few countries (eTable 1 in the Supplement). Herein, we use a study-replication design with data from 5 large and independent samples, addressing generalizability of results and the increasing concern of bias and nonreproducibility of results from research studies.\textsuperscript{36,35}

Limitations

Our study has several limitations. Despite its large overall sample size, the effective sample size for individual countries was limited by the low prevalence of ASD. Misspecification is another potential limitation. The first potential misspecification arises from the possible violation of the assumption of independence between genetic and environment. If this correlation is not specifically included in the model, its components will mostly be incorporated into the estimate of genetic variance component, potentially biasing the heritability estimate. The direction of the bias will depend on the sign of the covariance between genetic and environmental factors.\textsuperscript{36} The second misspecification arises from plausible gene-environment interactions that were not modeled and could also bias the heritability estimate. The direction of bias will depend on whether the environmental component is familial and whether the trait is multifactorial.\textsuperscript{30} One potential interaction, or subgroup difference, is the difference in health care-seeking behavior between sister and sister and brother and brother. For example, sisters might be more likely to share information and encourage early identification in their children. To adjust for this, the model should allow for differences in shared environment effects for different parental sibling types or include an interaction between shared environment and sibling type, which would require an even larger sample size. Furthermore, misspecification of C and M will cause an upward bias of the M component because the M effect is the only effect that is sibling-type specific (correlation of M is only present in the sister-sister pair and absent in other
sibling pairs). Nevertheless, because most of the M estimates are close to zero (Table 3), the risk of upward bias should be minimal. Any study using differences in ASD occurrence between sibling pairs will rely on the assumption of independent ascertainment. Lack of independence will make pairs more similar and may therefore inflate the contribution from the additive genetic effect as well as shared environment. Using differences in ASD variation between monozygotic and dizygotic twins could have strengthened the estimation of shared environment. However, because we did not have information about zygosity status and because twins are well known to have an elevated risk of being diagnosed with ASD, we excluded twins from our calculations.

Conclusions

Based on population data from 5 countries, in what is to our knowledge the largest study to date, the heritability of ASD was estimated to be approximately 80%, indicating that the variation in ASD occurrence in the population is mostly owing to inherited genetic influences, with no support for contribution from maternal effects. The results suggest possible modest differences in the sources of ASD risk between countries. The contributions of gene-environment interactions or correlations between genes and environment to ASD risk are important unanswered questions.
Heritability of autism spectrum disorders:


Most published research findings are false-but a little replication goes a long way. PLoS Med. 2007;4(2):e28. doi:10.1371/journal.pmed.0040028


Limitations of GCTA as a solution to the missing heritability problem. Proc Natl Acad Sci USA. 2016;113(1):E61-E70. doi:10.1073/pnas.1520199113