Am. J. Hum. Genet. 69:1062–1067, 2001

Excess of Twins among Affected Sibling Pairs with Autism: Implications for the Etiology of Autism

David A. Greenberg,[1,2] Susan E. Hodge,[1,2] Janice Sowinski,[3] and Doug Nicoll[3]

[1]Division of Statistical Genetics, Department of Biostatistics, Columbia University, and [2]New York State Psychiatric Institute, New York; and [3]Human Biological Data Interchange, Philadelphia

It is widely accepted that genes play a role in the etiology of autism. Evidence for this derives, in part, from twin data. However, despite converging evidence from gene-mapping studies, aspects of the genetic contribution remain obscure. In a sample of families selected because each had exactly two affected sibs, we observed a remarkably high proportion of affected twin pairs, both MZ and DZ. Of 166 affected sib pairs, 30 (12 MZ, 17 DZ, and 1 of unknown zygosity) were twin pairs. Deviation from expected values was statistically significant ($P < 10^{-6}$ for all twins); in a similarly ascertained sample of individuals with type I diabetes, there was no deviation from expected values. We demonstrate that to ascribe the excess of twins with autism solely to ascertainment bias would require very large ascertainment factors; for example, affected twin pairs would need to be, on average, ~10 times more likely to be ascertained than affected nontwin sib pairs (or 7 times more likely if "stoppage" plays a role). Either risk factors (related to twinning or to fetal development) or other factors (genetic or nongenetic) in the parents may contribute to autism.

Introduction

It is widely accepted that autism [MIM 209850] is extensively influenced by genetic factors (Smalley et al. 1988; Bailey et al. 1995). However, despite converging evidence from gene-mapping studies (e.g., International Molecular Genetic Study of Autism Consortium 1998, 2001; Ashley-Koch et al. 1999; Barrett et al. 1999; Philippe et al. 1999; Risch et al. 1999; Auranen et al. 2000; Buxbaum et al. 2001), it remains unclear what other factors may play a role in the risk for autism. In particular, there is controversy over whether perinatal factors play a role in the etiology of autism—and, if so, whether that role is causal or ancillary (Steffenburg et al. 1989; Bailey et al. 1995; Trottier et al. 1999). Yet another possibility is that factors in the *mother*—such as autoimmunity (Comi et al. 1999) or bleeding during pregnancy (Torrey et al. 1975)—may affect the development and/or perinatal environment of the fetus; however, such factors have received less attention.

In an extensive data-collection initiative, the Autism Genetic Resource Exchange (AGRE) of the Cure Autism Now (CAN) Foundation is collecting genotype and clinical data on a large cohort of families ascertained through at least two siblings with autism or autism-related conditions (e.g., pervasive developmental delay [PDD] or Asperger syndrome) (Geschwind et al. 2001). These data are being collected under the supervision of the Human Biological Data Interchange (HBDI). During evaluation of these pedigrees, we noticed a striking and dramatic excess of twin pairs in the data set, compared with the highest population frequencies that reasonably could be expected: among affected sib pairs (ASPs), we observed a 4–5-fold increase for DZ twins and much more than a 10-fold increase for MZ twins. These increases were statistically significant.

This observation, if confirmed, would indicate that simply being a twin represents a risk factor for autism. This risk may originate in the environmental conditions in the womb, including competition for nutrients, obstetric complications, or factors in the mother that predispose to autism—whether these factors are intrauterine, perinatal, immune-related, or inherited factors related to pregnancy, or even in vitro fertilization (IVF), although IVF is associated with an increase only in DZ, not MZ multiple births.

In this article, we describe the striking increase in twin pairs among the ASPs in the study. We demonstrate that, if this increase is due to sampling issues rather than to biological factors, then the required ascertainment-bias factor must be very high.

Data

For genetic studies, the AGRE and the HBDI recruit families with at least two affected relatives—most of whom are affected sibs—with autism and/or either PDD

Received June 15, 2001; accepted for publication August 20, 2001; electronically published October 2, 2001.

Address for correspondence and reprints: Dr. David A. Greenberg, NYSPI, Unit 24, 1051 Riverside Drive, New York, NY 10032. E-mail: dag@shallot.salad.mssm.edu

0002-9297/2001/6905-0015$02.00

or Asperger syndrome. Families are recruited both via mailings requesting participation, sent to members of CAN, and via presentations given at meetings of autism support groups and the Autism Society of America. No families are recruited through physicians. Families with MZ twins are not particularly sought, since one of the purposes of the AGRE is to collect families for linkage studies, and MZ twins add no information for linkage.

After families have been identified, trained investigators go to the family's home and administer the Autism Diagnostic Interview (Le Couteur et al. 1989). Affected offspring are coded as having autism, PDD, or Asperger syndrome. Most of the twins have undergone genotyping, and their zygosity, originally reported by parents, has been confirmed in all cases but one.

Although families with more than two affected offspring or with two or more other affected relatives are included in the AGRE study, we limited the current analysis to families having exactly two affected offspring (i.e., having no affected trios, etc.), to simplify the calculations and the handling of ascertainment. Also, we also did not include families with "mixed twin pairs"—that is, families with one affected twin, one unaffected cotwin, and a nontwin affected sib.

We divided the ASPs into two diagnostic categories: "narrow" (i.e., both affected sibs have autism) and "broad" (i.e., one or both of the affected members have either PDD or Asperger syndrome). Table 1 shows the observed distribution of ASPs, by diagnosis, by twin status, and by sex distribution. Finally, we also examined data on similarly ascertained families from the HBDI study of type 1 diabetes, or insulin-dependent diabetes mellitus (IDDM), for comparison with the data on autism.

Statistical Analysis

Across populations, published twinning rates (per birth) vary widely for DZ twins but are fairly consistent for MZ twins. We used values from Thompson et al. (1991), who report, for European populations, rates of 1/125, for DZ twins, and 1/260, for MZ twins, because these values are at the high end of the ranges cited by Vogel and Motulsky (1979). Expected twinning rates per sib pair are approximately double the per-birth rates. (If r_{birth} is the twinning rate per birth, and r_{pair} is the rate per sib pair, then $r_{\text{pair}} = 2r_{\text{birth}}/(1 + r_{\text{birth}}) \approx 2r_{\text{birth}}$, since r_{birth} is small.) Thus, the expected proportions of sib pairs who are twin pairs are .016 (2/125) for DZ twins, .008 (2/260) for MZ twins, and .024 for all twins.

We performed the statistical analyses on the narrow and the total (i.e., narrow + broad) diagnostic groups. The observed twinning rates in our sample showed substantial and statistically significant deviations from population rates (table 2).

Table 1

Distribution of ASPs

Group	No. in Diagnostic Category[a]		
	Narrow	Broad	Total
Singletons	82 (47,32,3)	54 (28,21,5)	136 (75,53,8)
DZ twin	8 (6,2,0)	4 (2,2,0)	12 (8,4,0)
MZ twin	15 (14,—,1)	2 (1,—,1)	17 (15,—,2)
UZ twins	1 (1,0,0)	0 (0,0,0)	1 (1,0,0)
Total	106 (68,34,4)	60 (31,23,6)	166 (99,57,10)

[a] Numbers in parentheses indicate breakdown into male-male, mixed sex (if applicable), and female-female, respectively. UZ = unknown zygosity.

Could these results be explained by preferential ascertainment of twin ASPs over nontwin ASPs? To examine this question, we define the "ascertainment factor," *A*, as the ratio of the probability that a twin ASP will be ascertained versus the probability that a nontwin ASP will be ascertained (see table 2). To explain our observations, *A* would have to equal, for example, 11.9 for all twins (narrow diagnosis). This means that these twin pairs would have to be 11.9 times more likely to be ascertained than nontwin ASPs, in order to match the excess of twins that we observed. Moreover, the lower 95% confidence limit (95%CL) on that *A* is 7.2; that is, 7.2 is the lowest possible value of the ascertainment factor that could render our observations "not significant" at the .05 significance level (two sided). The results for MZ twins (narrow diagnosis) are even more striking: *A* would have to be >22, with a lower limit of almost 12.

For comparison, we also examined proportions of twin pairs among ASPs in HBDI's study of IDDM that have been collected during the past ~20 years, in a manner similar to that used for the collection of data on autism. Using the same criterion as was used for the families with autism—that is, exactly two affected offspring—we identified 649 ASPs, 15 (13 MZ, 1 DZ, and 1 of unknown zygosity) of whom were twins. Compared to population rates, these observations represent a statistically significant deficit ($P < .001$) for DZ twins and a statistically significant excess ($P < .01$) for MZ twins; but for all twins there is no significant difference.

Discussion

We have documented a striking increase in the proportion of twins among ASPs with autism, an excess not observed in a similarly ascertained population of ASPs with IDDM. We have also quantified the magnitude of ascertainment bias that would be required to explain this increase, if the increase is due to sampling issues, without a biological basis. For example, twin ASPs overall would have to be ~9–12 times more likely to be

Table 2

Deviations, from Population Rates, of Observed Proportions of Twins—and Values of *A* That Are Needed to Explain Observations

		Narrow Diagnosis			Narrow + Broad Diagnoses		
Twin Group	Population Rate	Rate Observed	P[a]	A[b] (Lower 95%CL)	Rate Observed	P[a]	A[b] (Lower 95%CL)
DZ	.016	.075 (8/106)	<.001	5.9 (2.5)	.072 (12/166)	<.00005	5.4 (2.7)
MZ	.008	.142 (15/106)	<.000001	22.4 (11.8)	.102 (17/166)	<.000001	15.2 (8.7)
All[c]	.024	.226 (24/106)	<.000001	11.9 (7.2)	.181 (30/166)	<.000001	9.0 (5.8)

[a] Two-sided, exact binomial calculations.

[b] $A = (r_{observed}/s_{observed})/(r_{population}/s_{population})$, where r denotes rate of twin pairs (of specified type) among sib pairs and s denotes rate of nontwin pairs among sib pairs.

[c] Includes UZ twins, in addition to MZ and DZ twins.

ascertained than nontwin ASPs (or somewhat less if "stoppage" plays a role; see below).

Furthermore, we did not include in the calculations (*a*) two sibships, each with a twin ASP plus an affected nontwin; (*b*) six "mixed twin pairs"; (*c*) three sets of triplets, each with two affected members; and (*d*) one set of quadruplets, all affected. The existence of these sib groups, together with the twin data, suggest that multiple birth is an important risk factor for autism.

Several questions concerning our findings arise: What conditions must prevail if our findings are to be explained solely by ascertainment bias? How appropriate a comparison sample is provided by the IDDM twin data? What implications do our findings have for understanding the etiology of autism?

Ascertainment Bias

If ascertainment does explain the excess of twins in this data set, how could such an ascertainment difference arise? We consider several possibilities:

1. Parents of twins with autism may be much more likely to volunteer for a genetic research study than are parents of *two nontwin* children with autism. This explanation might hold if (*a*) the burden of having two children whose autism emerges simultaneously is much greater than the burden of having two children whose autism emerges ⩾1 year apart and (*b*) the probability of parents' volunteering to participate in a genetic study is proportional to this burden. These two factors may hold but would have to be demonstrated.

2. Parents could be more likely to perceive autism as being genetic when twins are affected than when nontwin siblings are affected—and that this perception predisposes them to participate in a genetic study.

3. MZ twins may have been turned away by other genetic studies and therefore may participate disproportionately in the AGRE study. Many genetic studies of autism currently concentrate on *linkage* analyses, to which MZ twin pairs contribute no information. The AGRE, in contrast, collects twin data and also identifies families for genetic linkage studies, although (as mentioned above) it does not preferentially collect twins. Thus, conceivably, if MZ twins' parents who want to contribute to research have been turned away by other studies, such twins may emerge in increased proportions in the AGRE data set; however, this would not explain the dramatic increase of DZ twins in the AGRE data set, since DZ twins are just as acceptable for linkage studies as are nontwin sib pairs.

4. The ascertainment differences may be the result of "stoppage" (Jones and Szatmari 1988; Slager et al. 2001). In the extreme situation of "complete" stoppage, in which parents always stop having children after the birth of their first affected child, the *only* families to have an ASP would be those with an affected twin pair (or affected triplets, etc.) Although that is clearly not the case in our data set, since it does not contain many nontwin ASPs, stoppage still is a reasonable factor to consider. The phenomenon of stoppage may be viewed as a particular form of ascertainment bias, since, if stoppage holds, some proportion, d, families who "would have had" a second affected child fail to do so because they have stopped having children after the first affected child. Other calculations that we have done on the complete data set (authors' unpublished data) indicate that the stoppage parameter, d, may be as high as 30%. Incorporating d into our calculations would attenuate the magnitude of the ascertainment factor A by a factor $1 - d$, or ≈.7. Thus, for all twins, A would become 8.3 and 6.3, for the "narrow" and "total" diagnostic groups, respectively, with lower 95%CLs of 5.0 and 4.1, respectively. These reduced A values are still quite high.

Another factor arguing for twinness as a risk factor for autism was the presence of a surprisingly large number (6/172) of families in which discordant twins were present but in which a third, nontwin child had autism and/or PDD. We did not include these in our calculations, because of assumptions that we would have had

to make about the ascertainment mechanism. Nonetheless, these families provide additional evidence that being a twin may be a risk factor for autism and/or PDD.

Data on Twins with IDDM

Another issue is whether the data on twins with IDDM provide an appropriate comparison with data on twins with autism. The families with IDDM were recruited via a mechanism similar to that used to recruit the families with autism and, basically, via the same organization. It is true that autism represents a much greater burden for parents than does IDDM. However, the issue is not whether twin ASPs with autism are more likely to be ascertained than are twin ASPs with IDDM but whether the increase in ascertainment of twin ASPs with autism over nontwin ASPs with autism is greater than the increase in ascertainment of twin ASPs with IDDM over nontwin ASPs with IDDM. This could be the case, but we know of no evidence for it.

The twinning rates among the ASPs with IDDM presented a less clearcut picture than did those among the ASPs with autism. Compared with what would be expected on the basis of population rates, the rate for DZ twins was lower but the rate for MZ twins was higher. Moreover, the overall twinning rate among the ASPs with IDDM was exactly as expected. Had the ASPs with IDDM shown the same rate of twinning as did the ASPs with autism, we would have expected to observe either 117 or 147 twin pairs overall (depending on whether we used, respectively, the "narrow" or "total" rates in table 2), among the 649 ASPs with IDDM—instead of the 15 that we did observe. One might expect that the number of twins in the IDDM data set would be greater than what would be predicted on the basis of the population figures, if presence of a disease were a factor affecting participation in the study. The fact that we did not observe the same kind of excess of twinning in another sample of ASPs with a serious disease, who were ascertained via a mechanism similar to that used for the sample with autism, provides some evidence, although it is not conclusive, against ascertainment bias alone as the explanation for the excess of twinning in the autism data set.

Implications for Etiology of Autism

The observation of such a notable excess of multiple births suggests several hypotheses. First, MZ twins are more overrepresented than are DZ twins, thus supporting the hypothesis that genetic factors do play a role in the development of autism. However, if being a twin is an important risk factor for autism and/or PDD, then one must ask whether the difference in the MZ:DZ ratio in autism is due to differences in the twinning processes for MZ twins versus DZ twins or to the intrauterine environment. For example, DZ twins develop from different ova, whereas MZ twins develop from the same ovum. Being monochorionic or monoamniotic may play a role in the competition for intrauterine resources, since monochorionic twins may exhibit significantly more perinatal mortality (Chitrit et al. 1999) and/or morbidity (Minakami et al. 1999) than do dichorionic twins. Also, since MZ twins result from a splitting of a single ovum, the reduction in the mass of the ovum may play a role. If, as our study suggests, being a twin is a risk factor for autism, then estimates, in the literature, of the genetic contribution to autism that are derived from the MZ:DZ ratio may be biased upward.

Second, if the magnitude of the MZ:DZ ratio is in fact due to genetic factors, then the observation of higher-than-expected proportions of twins, both MZ and DZ, suggests that those genetic factors may be related to genes in the parents rather than to genes in the offspring. In that case, linkage studies would need to focus on parents of children with autism and/or PDD and on the parents' sibs. In such a case, the multiplex families favored for linkage studies should consist of adult siblings with affected offspring, rather than the affected offspring themselves. Genetic effects may involve the intrauterine environment, in which case either the mother of the offspring with autism and/or PDD would be the "affected" member or some unknown interaction between the expression of paternal genes in the offspring and the mother's genetic complement could be a factor (e.g., as in Rh incompatibility).

Third, there is undoubtedly heterogeneity within the diagnosis of autism and/or PDD. The observation of excess twins among offspring with autism may indicate that the etiology in a notable proportion of offspring with autism is related to intrauterine or perinatal factors. On the other hand, the proportion of children with autism for whom the putative maternal or perinatal effects are primary may still be small, if those factors merely modify direct genetic effects. In either case, it becomes essential to attempt to differentiate between perinatally related etiologies and genetically related etiologies.

Finally, if the phenomenon described here is real, then why has no one else reported an excess of twins among families with autism? The only studies related to this question have been twin studies (i.e., studies whose entire sample consisted of twin pairs). The investigators compared the size of the total sample to the total "expected" number of affected twin pairs in the population of the appropriate country. However, this approach seems susceptible to error, since investigators may not have ascertained all the affected twin pairs in that country. One of these studies (Gillberg and Steffenburg 1990) reported a slight but nonsignificant excess of affected twin pairs over what would be expected, and the other did not (Bailey et al. 1995). To our knowledge, no one

else has examined the number of twin pairs *among* a sample of ASPs. Studies such as those by Philippe et al. (1999) and Risch et al. (1999) do not record whether their ASP samples included twin pairs—and, if so, how many (although Risch et al. mention two twin ASPs whose zygosity was questioned). We encourage other investigators of autism to see whether our intriguing finding is replicated in their data sets.

Acknowledgments

This work was supported in part by National Institutes of Health grants NS27941, DK31775, DK31813, and MH48858. We thank the HBDI and the AGRE for access to twin data from the autism and IDDM databases. We gratefully acknowledge the resources provided by the AGRE consortium and the participating AGRE families. The members of the AGRE consortium are as follows: AGRE Scientific Advisory Board—chairman Daniel H. Geschwind (University of California, Los Angeles [UCLA]) and members W. Ted Brown (New York State Institute for Basic Research in Developmental Disabilities, Long Island), Maya Bucan (University of Pennsylvania, Philadelphia), Joseph Buxbaum (Mt. Sinai School of Medicine, New York), Edwin H. Cook (University of Chicago, Chicago), T. Conrad Gilliam (Columbia Genome Center, New York), David A. Greenberg (Mt. Sinai Medical Center, New York), David H. Ledbetter (University of Chicago, Chicago), Bruce Miller (University of California, San Francisco), Stanley F. Nelson (UCLA School of Medicine, Los Angeles), Jonathan Pevsner (Kennedy Krieger Institute, Baltimore), Jerome I. Rotter (Cedars-Sinai Medical Center, Los Angeles), Gerard D. Schellenberg (University of Washington, Seattle), Carol A. Sprouse (Children's National Medical Center, Baltimore), Rudolph E. Tanzi (Massachusetts General Hospital, Boston), and Kirk C. Wilhelmson (University of California, San Francisco); AGRE Collaborator—Jeremy M. Silverman (Mt. Sinai Medical School, New York).

Electronic-Database Information

The accession number and URL for data in this article are as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim (for autism [MIM 209850])

References

Ashley-Koch A, Wolpert CM, Menold MM, Zaeem L, Basu S, Donnelly SL, Ravan SA, Powell CM, Qumsiyeh MB, Aylsworth AS, Vance JM, Gilbert JR, Wright HH, Abramson RK, DeLong GR, Cuccaro ML, Pericak-Vance MA (1999) Genetic studies of autistic disorder and chromosome 7. Genomics 61: 227–236

Auranen M, Nieminen T, Majuri S, Vanhala R, Peltonen L, Järvelä I (2000) Analysis of autism susceptibility gene loci on chromosomes 1p, 4p, 6q, 7q, 13q, 15q, 16q, 17q, 19q and 22q in Finnish multiplex families. Mol Psychiatry 5:320–322

Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M (1995) Autism as a strongly genetic disorder: evidence from a British twin study. Psychol Med 25:63–77

Barrett S, Beck JC, Bernier R, Bisson E, Braun TA, Casavant TL, Childress D, et al (1999) An autosomal genomic screen for autism: collaborative linkage study of autism. Am J Med Genet 88:609–615

Buxbaum JD, Silverman JM, Smith CJ, Kilifarski M, Reichert J, Hollander E, Lawlor BA, Fitzgerald M, Greenberg DA, Davis KL (2001) Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity. Am J Hum Genet 68:1514–1520

Chitrit Y, Filidori M, Pons JC, Duyme M, Papiernik E (1999) Perinatal mortality in twin pregnancies: a 3-year analysis in Seine Saint-Denis (France). Eur J Obstet Gynecol Reprod Biol 86:23–28

Comi AM, Zimmerman AW, Frye VH, Law PA, Peeden JN (1999) Familial clustering of autoimmune disorders and evaluation of medical risk factors in autism. J Child Neurol 14:388–394

Geschwind DH, Sowinski J, Lord C, Iversen P, Shestack J, Jones P, Ducat L, Spence SJ, AGRE Steering Committee (2001) The Autism Genetic Resource Exchange: a resource for the study of autism and related neuropsychiatric conditions. Am J Hum Genet 69:463–466

Gillberg C, Steffenburg S (1990) Reduced optimality in the pre-, peri-, and neonatal periods is not equivalent to severe peri- or neonatal risk: a rejoinder to Goodman's technical note. J Child Psychol Psychiatry 31:813–815

International Molecular Genetic Study of Autism Consortium (1998) A full genome screen for autism with evidence for linkage to a region on chromosome 7q. Hum Mol Genet 7: 571–578

——— (2001) Genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. Am J Hum Genet 69:570–581

Jones MB, Szatmari P (1988) Stoppage rules and genetic studies of autism. J Autism Dev Disord 18:31–40

Le Couteur A, Rutter M, Lord C, Rios P, Robertson S, Holdgrafer M, McLennan J (1989) Autism diagnostic interview: a standardized investigator-based instrument. J Autism Dev Disord 19:363–387

Minakami H, Honma Y, Matsubara S, Uchida A, Shiraishi H, Sato I (1999) Effects of placental chorionicity on outcome in twin pregnancies: a cohort study. J Reprod Med 44:595–600

Philippe A, Martinez M, Guilloud-Bataille M, Gillberg C, Rastam M, Sponheim E, Coleman M, et al (1999) Genomewide scan for autism susceptibility genes. Hum Mol Genet 8:805–812

Risch N, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J, Kalaydjieva L, et al (1999) A genomic screen of autism: evidence for a multilocus etiology. Am J Hum Genet 65: 493–507

Slager S, Foroud T, Haghighi F, Spence MA, Hodge SE (2001) Stoppage: an issue for segregation analysis. Genet Epidemiol 20:328–339

Smalley SL, Asarnow RF, Spence MA (1988) Autism and genetics: a decade of research. Arch Gen Psychiatry 45:953–961

Steffenburg S, Gillberg C, Hellgren L, Andersson L, Gillberg IC, Jakobsson G, Bohman M (1989) A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden. J Child Psychol Psychiatry 30:405–416

Thompson MN, McInnes RR, Willard HF (1991) Thompson and Thompson genetics in medicine, 5th ed. WB Saunders/ Harcourt Health Sciences, Philadelphia

Torrey EF, Hersh SP, McCabe KD (1975) Early childhood psychosis and bleeding during pregnancy: a prospective study of gravid women and their offspring. J Autism Child Schizophr 5:287–297

Trottier G, Srivastava L, Walker CD (1999) Etiology of infantile autism: a review of recent advances in genetic and neurobiological research. J Psychiatry Neurosci 24:103–115

Vogel F, Motulsky AG (1979) Human genetics: problems and approaches. Springer, New York