Heritable and Nonheritable Risk Factors for Autism Spectrum Disorders

Craig J. Newschaffer, Daniele Fallin, and Nora L. Lee

From the Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD.

Received for publication June 17, 2002; accepted for publication January 15, 2003.

INTRODUCTION

Autism spectrum disorders (ASDs) are developmental disabilities where language development is absent or delayed, rote or repetitive behaviors typically emerge, and nonverbal communication, imagination, and social interactions are profoundly hindered (1). The severity of impairment in each of these dimensions can be quite variable, as can individual cognitive functioning (2). However, even higher-functioning persons with ASD are confronted with significant lifelong challenges.

The first clinical descriptions of ASD were offered almost simultaneously by psychiatrist Leo Kanner and pediatrician Hans Asperger in the early 1940s. In his original paper, Kanner remarked on “the children’s aloneness from the beginning of life” (3, p. 250), implying that the constellation of behaviors that he observed was the result of pathology he believed to be present at birth. Despite this, because few autistic offspring had autistic parents, no obvious chromosomal anomalies were consistently found among autistic children, and the sibling recurrence rate was erroneously thought to be low (first estimated at only around 2 percent), geneticists initially doubted a role for heredity in ASD etiology (4). It was the publication of the first twin study on autistic disorder (5) that stimulated a stark change in the conventional wisdom. Findings of substantive differences in monozygotic-dizygotic twin concordance in the original and subsequent twin studies, coupled with higher estimates of sibling recurrence from studies conducted in the 1980s, indicated strongly that liability to ASD was heritable. Unfortunately, the subsequent two decades of genetic studies have failed to reveal a satisfactorily complete picture of ASD etiology.

This paper provides a review of ASD epidemiology, focusing first on what is known about heritable and nonheritable risk factors and concluding with a discussion of a more novel epidemiologic study design that blends elements of family-based and association studies in a way that may be necessary to explicate the risk factors underlying this major neurodevelopmental disorder.

DIAGNOSIS AND CASE DEFINITION

The behavioral symptom triad of impaired communication, impaired social interaction, and narrow range of interests has always been central to the diagnosis of ASD. In the Diagnostic and Statistical Manual of Mental Disorders, Fourth Revision (DSM-IV) (6), the category of “pervasive developmental disorders” includes the specific diagnoses of autistic disorder, essentially the phenotype originally described by Kanner, and Asperger’s syndrome, Rett’s syndrome, childhood disintegrative disorder, and pervasive developmental disorder not otherwise specified (6). The term ASD is now commonly used to refer to the pervasive developmental disorder diagnoses, sometimes exclusive of the etiologically distinct Rett’s syndrome (7) and developmentally distinct childhood disintegrative disorder (8). This term deemphasizes what many believe is an artifactual distinction among subtypes (9) and is consistent with the idea that different elements of the core triad of behavioral impairment may manifest in a myriad of different combinations. Much of the early epidemiologic and etiologic research focused on individuals with autistic disorder, because that has been the longest and best described phenotype.

The concept of “spectrum” also raises the possibility that the continuum of impairment may extend outside clinical bounds to those with combinations of milder social abnormalities and communication impairment and less rigid interest restrictions. On the basis of this possibility, several investigators have searched for a “broad autism phenotype” among relatives of clinical ASD cases. A recent review found that these studies consistently report higher rates of...
milder impairments (based on various definitions including social aloofness, lack of friendships, impaired play, and so on) in parents and, to a lesser extent, in siblings of probands with ASD than in controls (10).

**PATHOPHYSIOLOGY**

Although ASD is acknowledged as a brain pathology, no single distinguishing neuropathologic feature has yet been identified, and no single model of pathophysiology is currently accepted. Numerous physiologic abnormalities have been reported, and only those that have been most widely discussed are reviewed here.

Macroecephaly, large head size, was one of the phenotypic features of autistic disorder originally described by Kanner (3). Since then, this observation has been replicated in several, but not all, clinical autism series. One author reviewing this literature recently concluded that roughly 20 percent of children with autistic disorder across studies had head circumferences 2 or more standard deviations above average (11). A recent study compared ASD prevalence in population-based samples with and without previously documented infantile macrocephaly, finding fivefold greater prevalence in the macrocephaly group, but the total number of cases here was small (12). There have also been consistent findings that children with macrocephaly at the time of diagnosis tend to have normal head size at birth (13, 14). Hence, it may be reasonable to posit that the pathophysiologic mechanism leading to ASD could also be associated with postnatal changes in head size.

Data from autopsy and neuroimaging studies, although all based on fairly small samples, have generally supported the idea that the brains of individuals with ASD, at least in childhood, may be larger and heavier than average. One recent investigation looking at both brain volume and head size in subjects of various ages found adolescent and adult brain volumes in autistic subjects to be comparable with those of controls, while head size was larger (15). This suggests that brain volume may have been atypically large earlier in life.

Imaging studies have also given indications of particular brain regions showing enlargement or other anomalies. Autopsy investigations have generally found either subcortical forebrain anomalies of the limbic system or anomalies of the cerebellum in patients with autistic disorder (16, 17). Neuroimaging findings have suggested that increased brain volume in persons with autistic disorder is predominately a function of increased white matter volume (18). Abnormalities in brainstem nuclei have been replicated across autopsy and imaging studies (16, 19), as have findings of reduced numbers of Purkinje cells in the cerebellum. Neuroanatomic studies have recently established simultaneous occurrence of cerebellar and frontal lobe defects (17, 20). Among the prominent autopsy findings in the limbic system are abnormalities in the amygdala. There has been considerable interest in this structure because of existing evidence implicating a role for the amygdala in primate social behavior, reports of “acquired autism” following amygdalotomy, and neuroimaging findings, including some functional studies, documenting amygdalar anomalies in persons with ASD (21). Recent functional imaging studies of persons with ASD have called particular attention to both the amygdala and the fusiform gyrus (22), a structure apparently important in the recognition of faces, but more replication is needed.

In addition to neuroanatomic studies, serologic studies have generated hypotheses about ASD pathogenesis. For example, some involvement of the serotonergic system appears likely. In addition to its role as a neurotransmitter, serotonin is believed important in regulating neuronal differentiation, synaptogenesis, and neuronal migration during development (23). Schain and Freedman (24) first reported elevated whole blood serotonin in individuals with autistic disorder compared with controls, a finding replicated in a number of reports (23). The serotonergic hypothesis has also been supported by clinical studies showing some success of serotonin reuptake inhibitors in moderating certain inappropriate behavior among subjects with autistic disorder (25–27). Positron emission tomography scan studies have documented that children with autistic disorder do not exhibit the typical age-associated changes in serotonin synthesis capacity seen in nonautistic children (28).

Serologic studies have also pointed to possible links with the immune system. Subjects with ASD were more likely than controls to have detectable levels of central nervous system autoantibodies in sera (29, 30), but these have not yet been correlated with specific pathophysiologic effects. In the case of autoantibodies to myelin basic protein, signs of the expected effect, demyelination, have not been observed in persons with ASD (31). Other immune system irregularities, including decreased function of natural killer and T cells as well as decreased immunoglobulins (32), have also been found with increased frequency in subjects with ASD compared with controls.

One recent serologic study (33) has spurred much interest in a possible new class of ASD biomarkers. Levels of nine neuropeptides or neurotrophins were measured in neonatal blood samples retained on heel-stick cards from small samples of children subsequently diagnosed with autistic disorder, mental retardation (no ASD), or cerebral palsy as well as a group of controls. There were several proteins where almost all the children in both the autistic disorder and retardation groups had high levels, while the cerebral palsy and control groups had consistently low levels. The similar profile in both the mental retardation and autistic disorder groups is suggestive of a biomarker for nonspecific developmental dysfunction.

Overall, the large variety of neuropathologic changes noted and the variability seen across subjects imply that ASD is etiologically heterogeneous. Additionally, a number of the major anomalies observed, including cerebellar and brainstem findings, are very likely prenatal in origin. This, coupled with the recent findings of atypical neuropeptide profiles at birth, indicates strongly that the neuropathologic process underlying ASD begins in utero. Yet, brain plasticity may still allow for postnatal factors to affect the disease’s natural history.
DESCRIPTIVE EPIDEMIOLOGY

Prevalence

Two extensive reviews of the existing population-based ASD prevalence studies worldwide have appeared in the published literature within the last 2 years (2, 34). Both reviews acknowledge the extensive heterogeneity across studies. One concluded that the best prevalence for autistic disorder is nearly 5/10,000 (2), while the other offered 10/10,000 as a best estimate (34). The authors of both reviews concur that the prevalence of all ASD is manifold, from two to five times higher than that for autistic disorder alone.

More recent research studies have yielded prevalence estimates for autistic disorder many times higher than the 5–10/10,000 range (2). At this point, the heterogeneity in study design, source population, and criteria for and methods of identifying cases temper the inferences about secular trends that can be made from these studies.

Administrative data, the routine information collected by service delivery programs, likely have contributed more than research studies to the rising public concern over ASD prevalence trends. However, the potential for case-to-case and year-to-year inconsistencies in the categorizations used in administrative data makes these data more susceptible to ascertainment biases than the research studies. For example, the numbers of children classified with autism by state special education departments across the country have increased approximately 25 percent per year since 1994 (35).

This trend is not, however, unique to the autism special education classification, as the category of “other health impairments,” which includes children with attention deficit hyperactivity disorder, has also experienced increases of similar magnitude (35). In California, the Department of Developmental Services recently published a widely cited report on the numbers of individuals with autism registered with that agency, documenting large annual increases in the numbers of persons with an autism classification (36). When population denominators were applied to these autism case counts, the prevalence in most recent years was still near what would be expected on the basis of epidemiologic research studies (2).

At this point, it is not possible to say that the available data, research and administrative, clearly support the hypothesis that the underlying risk of ASD has been increasing with time. Knowing the true pattern in underlying risk over time is of great interest, because short-term increases in true disease risk would support a role for nonheritable mechanisms in ASD etiology.

ASD high-risk groups

The only identifiable group known for certain to have substantively elevated ASD risk is siblings of affected individuals (37). Certain rare medical disorders, in particular tuberous sclerosis, fragile X, and epilepsy, are also believed to place individuals at moderately higher risk for ASD (38, 39). Because the absolute ASD prevalence among males is still fairly low, males cannot be considered at high risk for ASD but, for unknown reasons, ASD does occur from three to four times more often in males than in females (40).

ASD subtypes

In addition to the DSM-IV diagnostic subcategories, there are other variable phenotypic features associated with ASD. If, as suspected, ASD is an etiologic heterogenous condition, subclassification of cases by different phenotypic features might help to reveal etiologically distinct subgroups. Cognitive impairment is one trait commonly used to subtype individuals with ASD. Approximately 70 percent of individuals with autistic disorder are cognitively impaired, 40 percent severely (2). The gender ratio among cases moves toward 1:1 for those with more severe cognitive impairment (2). The proportion of persons with cognitive deficits among individuals with all ASDs is likely lower than 70 percent, but no good estimate is yet available.

Much interest has also been expressed in regressive autism as an ASD subtype potentially possessing a unique etiology. ASD symptoms must be present before the age of 3 years to meet DSM-IV diagnostic criteria, but among cases there is a subgroup whose development appears typical up to 15–19 months, after which language decays and social problems emerge. The size of this subgroup is unclear, estimates ranging from 15 to 40 percent of all children with ASD (41), and population-based data are lacking. The presence of a subgroup of cases in which symptoms emerge later in infancy does not in itself imply the existence of a nonheritable risk factor; for example, clinical manifestations of the genetic condition sickle cell anemia do not emerge until most of the newborn child’s fetal hemoglobin is replaced by the variant hemoglobin.

Other approaches suggested for possibly meaningful subtyping of ASD are generally based on associated physiologic abnormalities, including the presence of minor morphologic anomalies (42, 43), seizure disorders (44, 45), gastrointestinal tract symptoms (46–48), and sleep disturbances (49, 50).

HERITABLE RISK FACTORS

Evidence for heritability

Several lines of evidence support a heritable component to ASD etiology, although no particular ASD-predisposing gene has been confirmed to date. The studies supporting a genetic component to ASD are summarized in table 1.

Twin studies. Monozygotic or “identical” twins share all of their genes, while dizygotic or “fraternal” twins share only half of their genes, on average. Accordingly, increased disease concordance rates among monozygotic twins versus dizygotic twins can provide compelling evidence for a heritable component to disease etiology.

In the late 1970s, the first study of multiple twin pairs reported four of 11 monozygotic pairs (36 percent) concordant for autistic disorder compared with zero of 10 concordant dizygotic pairs (0 percent) (51), providing provocative evidence for heritability. Subsequent twin information from a University of California, Los Angeles, study reported 96 percent monozygotic concordance (of 23 pairs) versus 30 percent dizygotic concordance (of 17 pairs) (52), confirming heritability. A contemporaneous Scandinavian twin study also reported a high monozygotic concordance (90 percent)
with no observable dizygotic concordance (53). However, a more recent British twin study, including the initial pairs from Folstein and Rutter (51), found only 60 percent monozygotic concordance (among 25 pairs) versus 0 percent dizygotic concordance (among 20 pairs) (54). These recent twin data also generate high estimates of heritability but do not fit a particular pattern of inheritance. Moreover, heritability estimates from twin studies may be overstated, one potential reason being the association between zygosity and chorionicity. Dizygotic twins always have separate sets of fetal membranes, while two thirds of monozygotic twins share a chorionic membrane (55). Because the placenta is formed from chorionic tissue, these monozygotic twins will share a placenta while dizygotic twins, and dichorionic monozygotic twins, will have two placentas. Consequently, monozygotic and dizygotic twins may experience different prenatal environmental influences. Monochorionicity, as opposed to monozygosity, has been linked to a number of adverse perinatal outcomes in twins (56).

Taken together, the twin studies imply some influence of environmental factors in addition to heritable predisposition to autistic disorder (54). An additional notable finding from the British twin studies was the reported excess concordance of broader ASD phenotypes among monozygotic twins (51, 54). This raises the possibility that the heritable trait, and therefore the predisposing genes, may be a broader underlying characteristic rather than with any particular disorders as clinically defined.

**Familial aggregation studies.** The large heritability estimates found in twin studies are supported by evidence of familial aggregation of ASD in sibling and population-based studies. For example, genealogic information available for a Utah-based population study allowed estimates of kinship (relatedness) to be calculated among all members of the study. Autistic disorder cases in this Utah registry had an estimated kinship greater than 20 times the average kinship among nonautistic individuals of the same birth year, suggesting familial aggregation (57).

### TABLE 1. Evidence for a genetic component of autism etiology

<table>
<thead>
<tr>
<th>Author(s) and year (reference)</th>
<th>Heritability</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twin studies</td>
<td>Folstein and Rutter, 1977 (5)</td>
<td>MZ* &gt;&gt; DZ*</td>
</tr>
<tr>
<td>Ritvo et al., 1985 (52)</td>
<td>MZ &gt;&gt; DZ</td>
<td></td>
</tr>
<tr>
<td>Steffenburg et al., 1989 (53)</td>
<td>MZ &gt;&gt; DZ</td>
<td></td>
</tr>
<tr>
<td>Bailey et al., 1995 (54)</td>
<td>MZ &gt;&gt; DZ</td>
<td></td>
</tr>
<tr>
<td><strong>Familial aggregation</strong></td>
<td>Smalley et al., 1988 (58)</td>
<td>λs* &gt;&gt; 1</td>
</tr>
<tr>
<td>Jorde et al., 1991 (59)</td>
<td>λs &gt;&gt; 1</td>
<td></td>
</tr>
<tr>
<td>Ritvo et al., 1989 (37)</td>
<td>λs &gt;&gt; 1</td>
<td></td>
</tr>
<tr>
<td>Jorde et al., 1990 (57)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overlap with other genetic disorders</strong></td>
<td>Brown et al., 1986 (66)</td>
<td>Fragile X</td>
</tr>
<tr>
<td>Ritvo and Piven, 1991 (67)</td>
<td>Fragile X</td>
<td></td>
</tr>
<tr>
<td>Folstein and Rutter, 1988 (65)</td>
<td>Neurofibromatosis</td>
<td></td>
</tr>
<tr>
<td>Steffenburg et al., 1996 (63)</td>
<td>Prader-Willi</td>
<td></td>
</tr>
<tr>
<td>Smalley, 1998 (64)</td>
<td>Tuberous sclerosis</td>
<td></td>
</tr>
<tr>
<td><strong>Genetic models</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mendelian</td>
<td>Ritvo et al., 1985 (75)</td>
<td>Autosomal recessive</td>
</tr>
<tr>
<td>Petit et al., 1996 (80)</td>
<td>X linked</td>
<td></td>
</tr>
<tr>
<td>Multigene</td>
<td>Jorde et al., 1991 (59)</td>
<td>Multifactorial additive threshold</td>
</tr>
<tr>
<td>Pickles et al., 1995 (76)</td>
<td>Multigene epistatic</td>
<td></td>
</tr>
<tr>
<td>Risch et al., 1999 (77)</td>
<td>Multigene epistatic</td>
<td></td>
</tr>
<tr>
<td>Other models</td>
<td>Cook et al., 1997 (70)</td>
<td>Imprinting</td>
</tr>
<tr>
<td>Schroer et al., 1998 (71)</td>
<td>Imprinting</td>
<td></td>
</tr>
<tr>
<td>Wolpert et al., 2000 (78)</td>
<td>Imprinting</td>
<td></td>
</tr>
<tr>
<td>Skuse et al., 1997 (85)</td>
<td>X linked + imprinting</td>
<td></td>
</tr>
<tr>
<td>Skuse, 2000 (86)</td>
<td>X linked + imprinting</td>
<td></td>
</tr>
</tbody>
</table>

* MZ, % of concordant autistic monozygotic twins; DZ, % of concordant autistic dizygotic twins; λs, sibling relative risk. 
Several studies have shown an increased risk for ASD among siblings of cases. This “sibling relative risk” is estimated as the ratio of the risk for ASD among siblings of cases to the risk, or prevalence, in the general population. Estimates of the probability of autistic disorder among siblings of cases range from 2 percent to 6 percent (58, 59), although some estimates are as high as 7 percent for siblings of male cases and 14 percent for siblings of female cases (37). Comparing these estimates with the accepted autistic disorder population prevalence estimates at the time of these sibling studies (4–6/10,000) provides very large sibling relative risk estimates in the range of 30–150.

Although such large sibling relative risk estimates are highly compelling, they are dependent on the population prevalence estimates. It is not implausible to expect that case ascertainment among families with affected probands could be better than that in population prevalence studies, leading to some inflation of sibling relative risk estimates. Much like the twin studies, familial aggregation of broader ASD phenotypes has also been observed (10, 60–62), again raising important questions about etiologic heterogeneity and definitions of “affected.”

**Overlap with known genetic disorders.** A separate line of support for genetic predisposition to ASD is the overlap with known genetic disorders such as Prader-Willi/Angelman syndrome (63), tuberous sclerosis (64), neurofibromatosis (65), and fragile X (66, 67). Further, abnormalities on almost every chromosome have been associated with some form of ASD phenotype, most notably on chromosomes 7, 15, and X (68). The most commonly cited of these are deletions and duplications of the proximal arm of chromosome 15 (69–72). Breakpoints for chromosomal inversions resulting in ASD features often lie within fragile regions of chromosomes, leading to speculation about the possible role of regions of unstable DNA and submicroscopic chromosomal deletions (73, 74).

### Genetic models for ASD

Given the evidence for a genetic component to ASD etiology, focus has naturally turned to gene discovery. Such efforts would achieve maximum power if a correct genetic model could be specified and assumed for linkage and association analyses. Several studies have sought to identify a particular genetic model through formal segregation analyses or analogy to other known genetic disorders.

**Segregation analyses.** Following up on their previous twin concordance results, Ritvo et al. (75) found evidence for autosomal recessive inheritance among 46 multiplex families. However, a more comprehensive study of all identified autistic disorder cases born in Utah between 1965 and 1984 failed to support a recessive major gene model (59). In the 185 Utah families identified, segregation analysis offered the most evidence for combined polygenic (many additive genes) and environmental effects in the absence of a major gene effect. This suggested a multifactorial threshold model in which several etiologic factors, perhaps some heritable and some nonheritable, would be needed to reach a critical liability threshold resulting in ASD (59).

In contrast to an additive threshold model, Pickels et al. (76) suggested an epistatic model of 3–10 predisposing interactive genes with a small proportion of cases being caused by other, potentially environmental, factors based on latent-class modeling of an underlying autistic status according to 15 observable phenotypes. An epistatic model, including at least 15 genes, has also been suggested from sibling allele-sharing estimates across the genome in families collected by Stanford University (77). Multigene models, either polygenic or epistatic, are congruent with the aggregation of broad ASD features among family members of probands, possibly reflecting possession of only a few predisposing variants rather than the full complement necessary to evoke a diagnosis.

**Chromosomal abnormalities.** In addition to additive and epistatic multigene models, several investigators have suggested imprinting and sex-linked genetic mechanisms through analogy to known chromosomal abnormalities resulting in ASD features. For example, several chromosome 15 abnormalities resulting in features characteristic of ASD are inherited solely from mothers (70, 71, 78), raising the possibility of an imprinting mechanism in gene expression. This is supported by evidence that many regions of chromosomes X and 15 are known to be imprinted (79) and by findings of maternal inheritance (73). An imprinting mechanism could explain the lack of convincing support for a Mendelian major gene model.

The overlap with fragile X syndrome and the excess of male cases of ASD have led many to speculate on a recessive X-linked inheritance model (80). However, family studies are incompatible with this hypothesis, given observations of some male-to-male transmissions and the exclusion of the X chromosome in some linkage studies (81–84). Further, association studies of the fragile X region with ASD have not been decisive (80). Recently, an intriguing model of imprinted X-linked inheritance has been proposed to explain both of these observations and has the further appeal of consistency with the male predominance in ASD (85, 86).

### Identification of ASD genes

**Linkage studies.** To date, six genome scans searching for linkage to an ASD gene have been published. These include the following: 1) a full scan of 152 sibling pairs, predominantly British (84, 87); 2) an autosomal scan of 75 families from the United States (88, 89); 3) a full scan of 51 families of predominantly European origin (90); 4) a full scan of 90 families from the United States (77); 5) a scan of 10 regions among 17 Finnish families (91); and 6) a full scan of 110 families from the United States recruited through the AGRE Program (92). Table 2 summarizes the findings from these studies.

Each scan pursued parametric or model-free linkage analyses based primarily on sibling pairs affected with autistic disorder; see Gutknecht (93) for a review. Given the uncertainty about the underlying genetic model for ASD, most scan results have focused on model-free affected pair strategies that do not require an assumption of the mode of inheritance. Affected relative pair methods compare the observed allele sharing between two relatives with the expected
sharing for that type of relative pair according to Mendelian laws. Significant departures from expected sharing values can be taken as evidence for linkage to a putative ASD gene in the region of excess sharing. However, the criteria used to define “significant” results from genome scans, which include between 300 and 500 markers, are controversial (94). Although general simulation results have provided guidelines for significance thresholds, the correct threshold will be unique to each study as it is based on the number of markers, the informativity of each marker in the study population, and the number and type of individuals studied. Such ambiguity has led to confusion in comparing results across studies. Several current reviews of recent findings have been published (93, 95, 96).

The most promising ASD region is on chromosome 7q, where linkage to a gene has been observed in four different scans (84, 88, 90, 97), and this signal has intensified with the addition of further markers in the region (73, 97). In addition, a maternally inherited paracentric inversion in the linked region has been identified in two brothers with autistic disorder and in a daughter with language pathology (73). A 7:13 translocation involving this region has also been observed in an autistic disorder case (98). Finally, a gene in this region with a mutation responsible for speech and language disorder has recently been identified, suggesting an overlap in genetic etiology of these disorders (99–101). Yet, to date, no mutations in this or surrounding genes have been directly identified in ASD families.

### TABLE 2. Genetic risk factors for autism

<table>
<thead>
<tr>
<th>Region/gene</th>
<th>Author(s) and year (reference)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chromosomal regions likely to harbor a gene</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7q</td>
<td>IMGSAC, Maestrini et al., 1999 (105)</td>
<td>MMLS* = 3.2</td>
</tr>
<tr>
<td></td>
<td>Barrett et al., 1999 (89)</td>
<td>MMLS = 2.2</td>
</tr>
<tr>
<td></td>
<td>Philippe et al., 1999 (90)</td>
<td>MMLS = 0.42</td>
</tr>
<tr>
<td></td>
<td>Ashley-Koch et al., 1999 (73)</td>
<td>MMLS = 1.77</td>
</tr>
<tr>
<td></td>
<td>Vincent et al., 2000 (98)</td>
<td>Translocation</td>
</tr>
<tr>
<td>2q</td>
<td>Buxbaum et al., 2001 (102)</td>
<td>HLOD* = 1.96</td>
</tr>
<tr>
<td></td>
<td>IMGSAC, 2001 (84)</td>
<td>MMLS = 3.74</td>
</tr>
<tr>
<td>16p</td>
<td>IMGSAC, Maestrini et al., 1999 (105)</td>
<td>MMLS = 2.93</td>
</tr>
<tr>
<td></td>
<td>Liu et al., 2001 (92)</td>
<td>MMLS = 1.93†</td>
</tr>
</tbody>
</table>

**Candidate genes**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Author(s) and year (reference)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-HTT</td>
<td>Cook et al., 1997 (103)</td>
<td>Association, short allele</td>
</tr>
<tr>
<td></td>
<td>Klauck et al., 1997 (104)</td>
<td>Association, long allele</td>
</tr>
<tr>
<td></td>
<td>Maestrini et al., 1999 (105)</td>
<td>No association</td>
</tr>
<tr>
<td></td>
<td>Zhong et al., 1999 (106)</td>
<td>No association</td>
</tr>
<tr>
<td></td>
<td>Persico et al., 2000 (110)</td>
<td>No association</td>
</tr>
<tr>
<td></td>
<td>Tordjman et al., 2001 (107)</td>
<td>Association, by severity</td>
</tr>
<tr>
<td></td>
<td>Yirmiya et al., 2001 (108)</td>
<td>Association, long allele</td>
</tr>
<tr>
<td></td>
<td>Betancur et al., 2002 (109)</td>
<td>No association</td>
</tr>
<tr>
<td>HOXA1</td>
<td>Ingram et al., 2000 (120)</td>
<td>Association</td>
</tr>
<tr>
<td></td>
<td>Li et al., 2002 (119)</td>
<td>No association</td>
</tr>
<tr>
<td>RELN</td>
<td>Persico et al., 2001 (118)</td>
<td>Association</td>
</tr>
<tr>
<td>HLA</td>
<td>Stubbs et al., 1985 (111)</td>
<td>Association</td>
</tr>
<tr>
<td></td>
<td>Daniels et al., 1995 (198)</td>
<td>Association</td>
</tr>
<tr>
<td></td>
<td>Warren et al., 1996 (113)</td>
<td>Association</td>
</tr>
<tr>
<td></td>
<td>Rogers et al., 1999 (114)</td>
<td>No linkage</td>
</tr>
<tr>
<td>GABRB3</td>
<td>Cook et al., 1998 (115)</td>
<td>Association</td>
</tr>
<tr>
<td></td>
<td>Maestrini et al., 1999 (105)</td>
<td>No association</td>
</tr>
<tr>
<td></td>
<td>Salmon et al., 1999 (116)</td>
<td>No linkage, no association</td>
</tr>
<tr>
<td></td>
<td>Martin et al., 2000 (117)</td>
<td>Association with nearby marker</td>
</tr>
<tr>
<td>HRAS</td>
<td>Herault et al., 1993 (123)</td>
<td>Association</td>
</tr>
<tr>
<td></td>
<td>Comings et al., 1996 (125)</td>
<td>Association</td>
</tr>
</tbody>
</table>

* IMGSAC, International Molecular Genetic Study of Autism Consortium; MMLS, multipoint maximum log of the odds (lod) score; HLOD, heterogeneity lod score. † Linkage in narrowly defined phenotype.
Other regions identified in multiple studies include 2q and 16p (84, 102). Some scans have detected a suggestion of linkage to 15q, potentially reflecting the region containing abnormalities associated with other genetic disorders exhibiting behavioral features similar to those of ASD, such as Prader-Willi (88, 90), although linkage has not been observed consistently in this region. Notably, none of the scans provided strong evidence for linkage to the X chromosome, despite correlations between ASD and fragile X syndrome and the predominance of males (43). However, these results do not preclude the possibility of an interaction between an X-linked gene and autosomal loci, which would be difficult to detect in the current reports.

Candidate genes. In addition to genome scans, several groups have pursued association studies of plausible candidate genes for evidence of polymorphisms that predispose to ASD. Efforts have focused on genes associated with biologic pathways implicated in ASD. Table 2 also includes a summary of this work.

Evidence supporting associations has been observed for the 5-hydroxytryptamine transporter (5-HTT) gene on chromosome 17 (103–110), the HLA-DR region (111–114), the gamma-aminobutyric acid A receptor b3 (GABRB3) gene on chromosome 15 (105, 115–117), the reelin (RELN) gene on chromosome 7 (118), the HOX genes (119, 120), fragile X genes (81, 121, 122), and the e-Harvey-ras (H Ras) gene on chromosome 11 (123–125). However, there has been no consistent replication of positive findings for any of these genes to date.

Inconsistencies in these genetic association studies may stem from reliance on the population genetic property of linkage disequilibrium to detect an association between what is actually a “marker” polymorphism in a candidate gene and the unobserved true ASD-predisposing variant. For example, the repeat polymorphism in the promoter region of the 5-HTT gene has been associated with ASD in several studies (103–110). However, studies performed in Germany and the United States observed association with different repeat sizes (103, 104, 108). A recent report suggests this may indicate an association between severity and repeat size (107). However, this more likely represents the differential linkage disequilibrium patterns in these distinct populations, such that the repeat marker alleles are associated with different underlying haplotypes in each population.

Interpretations

Evidence from twin studies, familial aggregation, and rare chromosomal abnormalities provide a compelling argument for some substantive heritable component in ASD etiology. However, no specific genes have been implicated.

The results of genome scans and candidate gene studies are difficult to interpret given the differences in populations, designs, and analytic techniques used. Meta-analysis may help to elucidate truly linked or associated regions across studies. However, these largely conflicting results more likely highlight the limitations of performing linkage and association studies in a complex and heterogeneous disorder. Etiologic mechanisms may include genetic variants in separate genes (locus heterogeneity), different variants within the same gene (allelic heterogeneity), and complicated epistasis and gene-environment interactions. One could consider a set of “etiologic classes” reflecting the different genes or interactive combinations resulting in ASD for particular subsets of individuals or families. The combination of families from different “classes” in the same linkage or association study will reduce the ability to detect the effects of any particular gene or “class.” Detection of a main effect will depend on the relative proportion of individuals carrying a particular genetic variant (or interactive combination that includes that gene) among the individuals studied. As this proportion is likely to fluctuate between data sets, it is unlikely that a particular linkage finding could be replicated in many other data sets, even if the same underlying model were at play.

Considering these complexities, gene identification studies would benefit most from better definition of phenotypes that correspond to a particular genetic etiology. For example, particular combinations of behavioral and/or physical symptoms may be indicative of a particular genetic predisposition. Restriction to that phenotype may decrease heterogeneity and allow a stronger detectable effect. Incorporation of particular models or interactive factors would also serve to reduce heterogeneity and focus on a particular etiologic class. The potential role of environmental influences, the observed overrepresentation of boys, and the potential for imprinting should be incorporated into genetic predisposition models and analyses. For example, if imprinting were important in the risk conferred by a particular gene, evaluating that gene without incorporating parental information could “wash out” the detectable effect. It is only through better characterization of phenotypes and inclusion of interactive factors or important covariates that genes underlying such a complex etiology will be discovered and consistently replicated.

NONHERITABLE RISK FACTORS

Recently, nonheritable ASD risk factors have again begun receiving attention (126, 127). In part, this is because a single, parsimonious model explaining ASD inheritance has not emerged. In addition, a small but compelling 1994 Swedish study (128, 129) reported a significantly greater than expected proportion of autistic disorder cases among members of a cohort prenatally exposed to thalidomide during days 20–24 of gestation, suggesting that exposure to an exogenous agent during a critical developmental period, in this case the time when the neural tube is formed, might cause ASD (130–132). Finally, the widely discussed idea that ASD prevalence may have risen markedly over the last decade has also promoted debate over nonheritable risk factors. However, given the evidence on heritability, it seems unlikely that truly sporadic cases of ASD would account for a substantial proportion of the disease in the population. Therefore, if environmental factors have a role in ASD etiology, they would most likely be important in mechanisms also involving some element of genetic susceptibility.
Obstetric suboptimality

As previously discussed, neurobiologic evidence points to prenatal initiation of pathophysiologic changes in the natural history of ASD. Given this and an absence of motivating hypotheses concerning particular exposures, one approach borrowed from other areas of perinatal epidemiology has been to look at summary measures of the “optimality” of the pregnancy and delivery (133–135). Optimality scales are based on the presence or absence of a myriad of factors that may be “suboptimal,” such as maternal age, maternal diabetes, neonatal respiratory distress (136), frequency of intercourse during pregnancy (138), placental insufficiency (139), and newborn slow to cry (140). Composite scores have been used in a number of epidemiologic studies (136, 138, 141–144), with most reporting lower optimality or higher composite risks among ASD cases than controls (136, 138, 142) (table 3).

However, aggregate measures of suboptimality, developed mainly to optimize small sample sizes, may not be the most appropriate means of examining prenatal and perinatal risk factors. Confounding by parity is one concern, since birth order effects are documented in ASD whether due to stoppage (145) or other phenomena (146). Adjusting for parity, however, has not consistently changed the association between ASD and suboptimality (140, 141, 146, 147).

Genetic predisposition may also be a confounding factor, because both ASD and obstetric suboptimality have familial components. Using unaffected siblings as controls (146) is only a crude means of controlling for family loading, and the possibility of residual confounding is supported by the observation of a positive association between the proportion of relatives affected with the broad autism phenotype and obstetric suboptimality in ASD probands (62, 147). Finally, the heterogeneity of suboptimality scores, typically a mix of antepartum, intrapartum, and postpartum factors, is a potential limitation. These varied factors are combined in a simple additive or multiplicative fashion to derive optimality scores that have been applied to perinatal mortality risk (135), but this model may not be appropriate to ASD etiology.

Specific prenatal factors

Three published case-control studies with primary data analyses of obstetric suboptimality included at least 50 in the case group and showed subanalyses for specific prenatal contributors to suboptimality (140, 148, 149). Two other published studies included over 50 cases but were secondary data analyses containing limited prenatal data (139, 150). Summaries of selected findings on nonheritable factors are included in table 3, and findings from studies that specific-

<table>
<thead>
<tr>
<th>Factor</th>
<th>Author(s) and year (reference)</th>
<th>Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suboptimality score</td>
<td>Finegan and Quarrington, 1979 (138)</td>
<td>Association*</td>
</tr>
<tr>
<td></td>
<td>Gillberg and Gillberg, 1983 (136)</td>
<td>Association*</td>
</tr>
<tr>
<td></td>
<td>Bryson et al., 1988 (142)</td>
<td>Association*</td>
</tr>
<tr>
<td></td>
<td>Lord et al., 1991 (141)</td>
<td>Association*</td>
</tr>
<tr>
<td></td>
<td>Piven et al., 1993 (146)</td>
<td>Association*</td>
</tr>
<tr>
<td></td>
<td>Cryan et al., 1996 (199)</td>
<td>Association*</td>
</tr>
<tr>
<td></td>
<td>Bolton et al., 1997 (147)</td>
<td>Association*</td>
</tr>
<tr>
<td>Maternal infection in pregnancy</td>
<td>Mason-Brothers et al., 1990 (148)</td>
<td>Flu/cold symptoms</td>
</tr>
<tr>
<td></td>
<td>Juul-Dam et al., 2001 (149)</td>
<td>Fever</td>
</tr>
<tr>
<td></td>
<td>Deykin and MacMahon, 1979 (151)</td>
<td>Exposure and clinical illness for a variety of infections</td>
</tr>
<tr>
<td>Prenatal or intrapartum medication</td>
<td>Deykin and MacMahon, 1979 (151)</td>
<td>Maternal self-report of any kind of medication use</td>
</tr>
<tr>
<td></td>
<td>Mason-Brothers et al., 1990 (148)</td>
<td>Obstetrics record review report of medication in any pregnancy</td>
</tr>
<tr>
<td></td>
<td>Fein et al., 1997 (166)</td>
<td>Labor induction during delivery</td>
</tr>
<tr>
<td>Parental preconception chemical exposure</td>
<td>Walker, 1976 (200)</td>
<td>Self-report</td>
</tr>
<tr>
<td></td>
<td>Felicetti, 1981 (168)</td>
<td>Self-report</td>
</tr>
<tr>
<td>Early childhood infection</td>
<td>Deykin and MacMahon, 1980 (140)</td>
<td>Self-report and medical records</td>
</tr>
<tr>
<td></td>
<td>IOM§ review, 2001 (183)</td>
<td>Expert review</td>
</tr>
<tr>
<td></td>
<td>MRC§ review, 2001 (41)</td>
<td>Expert review</td>
</tr>
</tbody>
</table>

* Association was statistically significant (p < 0.05).
† After adjustment for parity.
‡ For a number of infection measures.
§ MMR, measles-mumps-rubella; IOM, Institute of Medicine; MRC, Medical Research Council.
Associated with ASD. Chess (152) originally reported this association with pneumonia (151). The low frequency of reports suggests that infectious diseases known to affect the brain and ASD, including herpes simplex, rubella, syphilis, and varicella-zoster, have been composed mainly of occasional case reports (154). The published literature on other specific prenatal infections known to affect the brain and ASD, including measles, mumps, rubella, and influenza, has been impregnated with case reports (153). The published literature on other specific prenatal infections known to affect the brain and ASD, including measles, mumps, rubella, and influenza, has been impregnated mainly of occasional case reports (154). The low frequency of reports suggests that infectious diseases known to be associated with neuropathology are not a major independent cause of ASD (155).

Maternal infections. In the three studies, maternal infections were measured with nonspecific indicators, including maternal recall of fever and/or other symptoms and information archived in medical records. Each reported odds ratios above 1.0 that approached but did not attain statistical significance for the infection measure (140, 148, 149). Prior to their suboptimality paper, Deykin and MacMahon (151) published a more detailed analysis using the same sample of subjects but focusing on common viral diseases and infection markers during pregnancy. Data were derived from medical records or self-report of clinically diagnosed illness or illness exposure (defined as a case within the house). After adjusting for sibship size, odds ratios were significantly above 1.0 for chicken pox, herpes, and significantly above 1.0 for measles, mumps, rubella, and influenza and tended to be above 1.0 for chicken pox, herpes, and pneumonia (151).

Of specific infections known to affect the developing brain, rubella has been most commonly reported to be associated with ASD. Chess (152) originally reported this association, but upon further follow-up, six of the 18 original rubella-exposed cases were reclassified as not having ASD (153). The published literature on other specific prenatal infectious pathogens known to affect the brain and ASD, including herpes simplex, rubella, syphilis, and varicella-zoster, has been composed mainly of occasional case reports (154). The low frequency of reports suggests that infectious diseases known to be associated with neuropathology are not a major independent cause of ASD (155).

Prenatal and intrapartum pharmaceutical agents. As mentioned earlier, taking thalidomide during days 20–24 of gestation was clearly correlated with an increased risk for autistic disorder (128, 129), strongly suggesting that early prenatal xenobiotics could play a role in ASD etiology. Additional evidence comes from both animal studies (131, 132) and case series or reports (156–158) that prenatal use of valproic acid and other anticonvulsants also appears to increase the risk for ASD. Interestingly, the same drugs have had therapeutic benefit for nonepileptic children with ASD symptoms in a number of case reports (159–161). Because as many as 30 percent of ASD cases have comorbid epilepsy (40, 162, 163), there may be some overlap in etiologies (164).

Findings on other prenatal and intrapartum medications are mixed. Two of the three studies examined prenatal medication use, with one reporting a small excess in the ASD group for any medication use during pregnancy (140). One of the studies looking at labor induction found a significant association (149), but this was based on an external estimate of general population exposure. A large epidemiologic study in Japan found that the prevalence of autistic disorder among a 10-year birth cohort from one hospital that routinely used labor-inducing drugs was double that of the same years’ birth cohort from three other hospitals (165). The first hospital also had higher rates of use for general anesthesia, sedatives, and analgesics. In contrast, a recent report by Fein et al. (166) compared the frequency of labor induction in over 180 children with ASD with 197 language-impaired and 107 cognitively impaired controls. Each group had similar prevalences of induction, about 20 percent.

Preconception factors

The idea that preconception environmental exposures may be involved in ASD etiology arose in the 1970s from a retrospective case-control study of ASD that found a statistically significant difference in parental occupational exposure to chemicals during the preconception period (167). The retrospective nature of the self-reported exposures and the self-selection of the families studied cast doubt on the finding, but the finding was replicated in another smaller case-control study of 20 unselected families (168). In the 1990s, the hypothesis was revisited when unaffected parents who had lived near plastic manufacturing plants when they were young seemed to have more children with ASD (169). After investigations including a chromosomal study of the initial ASD cases and further case-finding efforts, the Massachusetts Department of Public Health concluded that further investigations were not warranted (170).

Beyond these studies of preconception chemical exposure and ASD risk, there seems to have been little interest in research on this topic, as reflected in the dearth of published epidemiologic studies. Interest in chemical exposures in the postnatal period has increased, however, on the heels of a potential ASD cluster in the New Jersey community of Brick Township. The findings of the Brick Township investigation are reviewed in the following section.

Postnatal factors

Chemical exposures. Hypotheses of postnatal chemical exposure and ASD have been investigated mainly through case studies and clinical series lacking comparison groups. The epidemiologic evidence for any specific postnatal environmental exposure leading to ASD is scant. One of the most comprehensive investigations took place in Brick Township, New Jersey, where the high local prevalence of ASD raised concern over possible connections to landfills in the area and possible drinking water contamination and/or chemical exposure through river swimming. The Agency for Toxic Substances and Disease Registry examined these possible exposure pathways, evaluating data on the levels of trihalomethanes, tetrachloroethylene, and trichloroethylene. Although the drinking water did contain contaminants at various points in time during the study period, the levels were either low or were in locations that did not correspond with the locations and timing of pregnancies of ASD cases (171, 172).

Infection. In addition to maternal prenatal infections and ASD, links between early childhood infections and ASD risk have also been explored. Several case studies have reported a sudden onset of autistic symptoms in older children after herpes encephalitis (173–175). Other infections that can result in secondary hydrocephalus, such as meningitis, have been implicated in ASD etiology, but even fewer such case reports can be found in the literature (154). Deykin and MacMahon’s (151) case-control study (controls were siblings) included maternal report and medical record docu-
mentation of exposure to, and clinical illness from, common viral illnesses in the first 18 months of life. After adjusting for sibship size, mumps, chickenpox, fever of unknown origin, and ear infections were all significantly associated with ASD risk.

**Measles-mumps-rubella vaccine.** Publication in 1998 of a paper reporting that eight of 12 children with regressive ASD referred to a pediatric gastroenterology department had measles-mumps-rubella vaccination prior to the onset of their developmental regression (46) catalyzed substantial public concern (176). Epidemiologic studies, however, have provided no evidence supporting a link between measles-mumps-rubella vaccination and ASD risk. The majority of the epidemiologic investigations have been ecologic comparisons, with consistent findings indicating no pattern of concomitant changes in measles-mumps-rubella coverage rates and ASD prevalence over time (177–181). Similarly, a case-only study comparing the report of symptom onset and diagnosis by time periods defined relative to measles-mumps-rubella vaccination date found no indication of post-vaccination increases in risk for ASD (177, 182). Comprehensive reviews of existing data on measles-mumps-rubella immunization and ASD, including reports by the Institute of Medicine (183), the Medical Research Council (41), and an expert panel convened by the American Academy of Pediatrics (184), concur that there is insufficient evidence to support measles-mumps-rubella as an ASD risk factor. Recently, a population-based, individual-level retrospective cohort study, including more than half a million children born in Denmark between 1991 and 1998 (82 percent of whom had received the measles-mumps-rubella vaccine), found no association between measles-mumps-rubella and ASD (185). The adjusted relative risks for autistic disorder and ASD, when compared with those of unvaccinated children, were 0.92 (95 percent confidence interval: 0.68, 1.24) and 0.83 (95 percent confidence interval: 0.65, 1.07), respectively.

**Mercury and thimerosal-containing vaccines.** Additional concern over vaccines and ASD stems from the use of thimerosal, a preservative containing ethylmercury. Ethylmercury is chemically similar to methylmercury, a known fetal neurotoxin that causes severe neurologic injury at higher doses and developmental delays and neurologic dysfunction at lower doses (186). Although limited animal and human data suggest that the toxicity of high-dose ethylmercury exposure is similar to that of high-dose methylmercury exposure (187, 188), the data on low-dose exposures to methylmercury are conflicting (189) and there are no data on low-dose exposures to ethylmercury. Bernard et al. (190) hypothesized that regressive ASD is related to mercury exposure, citing correspondence between traits and physiologic abnormalities of individuals with ASD and those with mercury poisoning and the fact that receipt of the recommended complement of childhood vaccines within the first 6 months of life could, depending on vaccine manufacturer and batch, exceed the US Environmental Protection Agency’s guideline for safe levels of methylmercury intake (186). In 1999, the American Academy of Pediatrics and the Public Health Service issued a joint statement recommending that vaccine manufacturers reduce or eliminate thimerosal in their vaccines (191). This prompted a similar statement by the US Food and Drug Administration (192).

Few epidemiologic data on this association have been assembled to date. Data from the Vaccine Safety Datalink of the Centers for Disease Control and Prevention suggest weak associations between thimerosal-related mercury exposure and any neurodevelopmental disorder, but not ASD specifically (193). However, there have been concerns over exposure and disease misclassification in these data (193). The recently completed Institute of Medicine expert review of thimerosal and ASD found the totality of existing evidence to be inconclusive (193). Further evaluation of the Vaccine Safety Datalink data has been proposed but, given the recent removal of thimerosal from vaccines and the difficulties involved in retrospective studies of vaccine-thimerosal exposure, further epidemiologic study of this risk factor will be challenging.

**Interpretations**

Currently, there is little evidence supporting any one nonheritable risk factor for ASD. However, the vast majority of existing studies have not been population based and have not been adequately sized to detect modest magnitude main effects. Given the likelihood of etiologic heterogeneity and genetic predisposition in ASD, it can be anticipated that nonheritable risk factors that are potentially quite important in certain subgroups (e.g., those with particular genetic predisposition) might not emerge as being significantly associated with ASD in small studies performed in select patient subgroups. Studies of obstetric suboptimality, in particular prenatal factors like infection and pharmacologic agent exposure, probably deserve further examination. In addition, as will be discussed further below, more consideration must be given to considering mechanisms by which these and other nonheritable factors may interact with susceptibility genes. Speculative risk factors, however, have received widespread media coverage within the last few years largely because of the strong degree of public concern, not the strong degree of existing evidence.

**ALTERNATIVE EPIDEMIOLOGIC APPROACHES FOR UNDERSTANDING ASD ETIOLOGY**

Clearly, existing epidemiologic and genetic research on ASD supports a complex etiology. The four main potential sources of risk for ASD are the following: 1) genetic predisposition of the mother, 2) environmental factors acting on the mother, 3) genetic predisposition of the child, and 4) environmental factors affecting the child. From these four sources, different etiologic models including heritable factors (single gene, additive, and epistatic), nonheritable factors, and their interactions can be developed (figure 1; table 4). Consequently, ASD research would likely benefit greatly from the use of study designs sufficiently flexible to detect all possible risk factors across a variety of etiologic models.

No existing studies have incorporated such parent-child complexities when testing for heritable or nonheritable ASD risk factors. For example, the affected sibling pair linkage
methods used in the various genome scans have been aimed at detecting genetic effects among probands (type C effects in table 4). In fact, the current affected sibling pair methods would simply not be able to detect maternal genetic effects in a powerful way, because the sisters of the mothers of those sibling pairs would be needed as the unit of analysis. In contrast, most epidemiologic studies have focused solely on the environmental effects acting upon the probands or their mothers (effect B or D in table 4). Ideally, a design and analysis strategy should be able to test for all the effects and their possible interactions in one setting. This framework has been proposed for the epidemiologic study of another neuropsychiatric disorder, schizophrenia (194), but to our knowledge has not been implemented there either.

Each of the commonly used genetic epidemiology designs can be extended beyond its typical use to evaluate additional effects, as noted in table 4, yet each current design falls short of complete flexibility. For example, one could stratify, or in some way condition, on the maternal genotype or environment when performing affected sibling pair linkage to detect type C effect and thus estimate $C \times A$, $C \times B$, or even $C \times D$ interactions. The main effects for these other factors, however, cannot be estimated in this setting. In the case-parent trio design often advocated for family-based association analyses, each family is conditioned on the child’s having ASD. Again, tests of interaction between the child’s genes and other nonheritable or heritable maternal factors could be performed, although the main effects of the factors could not be estimated. Further, tests of gene-environment interaction (and parent-of-origin effects) rely on the modest assumption of gene-environment independence (or no transmission distortion) within families, outside the risk for ASD (195). Studies of unrelated individuals, such as cohort or case-control designs, can estimate the main environmental effects, certainly of type D. Careful inquiry could also estimate the effects of type B, yet without genotyping the children or parents, the genetic effects of types A and C and their interactions with environmental factors cannot be evaluated.

If the genotypes of cohort or case-control participants and their parents are also collected, all the main effects could feasibly be addressed, as well as all the interactions, thus covering several plausible etiologic models in one setting and allowing the greatest flexibility by providing the opportunity for analyses within and across families to test specific hypotheses. We propose a case-parent/control-parent design for studies of the epidemiology of ASD as the most flexible and informative design to elucidate etiologic factors for this disorder.

The case-parent/control-parent design provides an opportunity to test genetic and environmental associations at the parent and child levels in several ways. Comparisons using the sampled cases and controls as the unit of analysis can certainly test each hypothesis, using logistic or conditional logistic (if controls are matched) regression approaches. In

<table>
<thead>
<tr>
<th>Design</th>
<th>Commonly tested effects</th>
<th>Potentially testable effects</th>
<th>Potentially testable interactions*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohort</td>
<td>$D$</td>
<td>$C, D$</td>
<td>$C \times D, C \times B, B \times D$</td>
</tr>
<tr>
<td>Case-control</td>
<td>$B, C, D$</td>
<td>$B, C, D$</td>
<td>$B \times C, B \times D, C \times D, B \times C \times D$</td>
</tr>
<tr>
<td>Affected sibling pairs</td>
<td>$C$</td>
<td>$C$</td>
<td>$C \times A, C \times B, C \times D$</td>
</tr>
<tr>
<td>Case nuclear families</td>
<td>$C$</td>
<td>$C, C^\dagger$</td>
<td>$C \times A, C \times B, C \times D$</td>
</tr>
<tr>
<td>Case-parent trios</td>
<td>$C$</td>
<td>$C, C^\dagger$</td>
<td>$C \times A, C \times B, C \times D$</td>
</tr>
<tr>
<td>Case-parent/control-parent</td>
<td>$A$</td>
<td>$B, C, C^\dagger, D$</td>
<td>All interactions</td>
</tr>
</tbody>
</table>

* $A$, effects of genes carried by mother; $B$, environmental effects in mother; $C$, effects of gene carried by autistic child; $D$, environmental effects in child.
† Can be tested via linkage and association in these designs.

Epidemiol Rev 2002;24:137–153
the world of emerging haplotype-based genetic analyses, parental genotypes can provide much more informative haplotype construction for these analyses. Within the same design, family-based tests can be performed among the case-parent trios to further elucidate parent-of-origin effects and provide tests of genetic linkage. The availability of control trios would then provide the opportunity to test assumptions such as the absence of general transmission distortion or gene-by-environment correlations.

From an epidemiologic perspective, the case-parent/control-parent design provides an opportunity to sample unrelated individuals (cases and controls) while taking advantage of the family-based genetic transmission information. In this way, the study sample does not overrepresent families heavily loaded with multiple affected members. If models of gene-environment interaction are important, samples not ascertained through a high genetic load may have a greater proportion of environmental etiology and thus be potentially more informative for complex etiologic models that include both heritable and nonheritable components. If the design can approach population-based sampling, then the study will have the further advantage of being able to estimate population exposure and allele frequencies as well as penetrances (risk estimates). Then, attributable risk estimates for particular genes or gene-environment combinations can also be constructed exclusively from study data.

The case-parent/control-parent design is still not without limitations. As with other approaches, etiologic heterogeneity can limit the design’s ability to reveal effects. In ASD, the emphasis has historically been on using the behavioral phenotype to create subgroups representing potentially distinct etiologic classes. Some have argued that focusing on groups tightly defined by behavior and excluding those with known ASD-associated comorbidities provide the most homogeneity for genetic purposes (196). Still others have argued quite the opposite, that the most fruitful studies will include a broad range of phenotypes, both with and without comorbidities, giving researchers maximum flexibility in exploring alternate subgroup classifications that may or may not correspond to existing clinical diagnostic categories (197). Implicit in the latter approach is the notion that existing means of categorizing the phenotype do not adequately reflect etiologic heterogeneity. Given a fixed sample size, the decision on whether to include narrow or broad phenotype cases will be related to researchers’ judgments on the anticipated size of effects to be detected and opinions on the likelihood that etiologic heterogeneity does or does not correlate with existing defined phenotypes.

Similarly, because no definitive ASD risk factors are known, researchers must also grapple with the question of how broad an array of potential risk factors to include in their studies. Of course, inclusion of data on a large number of potential risk factors increases the number of questions that can be investigated. In addition, should evidence accumulate supporting one particular factor, that variable could then be considered a potential confounder, and subsequent analyses could be conditioned on that factor, potentially improving the etiologic homogeneity of the groups under consideration. For example, once a gene is identified on chromosome 7, conditioning subsequent analyses on the child’s genotype at this predisposing locus may drastically improve the ability to identify subsequent genes, exposures, or interactive effects. The case-parent/control-parent design accommodates, to the extent that resources and respondent burden will bear, inclusion of heritable and nonheritable risk factors at both the parent and child levels, thus allowing consideration of a range of different etiologic models. Melding of the population- and family-based designs embodied in this approach could prove invaluable in advancing epidemiologic research on ASD.

ACKNOWLEDGMENTS

Dr. Newschaffer thanks the Cure Autism Now Foundation for its support of his review of autism prevalence studies. The authors acknowledge Laura Kresch Curran for her assistance in streamlining the manuscript.

REFERENCES

8. Dawson G. What is childhood disintegrative disorder, how is it different from autism, and what is believed to be its cause? J Autism Dev Disord 2002;30:177.

Epidemiol Rev 2002;24:137–153
55. Phillips DIW. Twin studies in medical research: can they tell us whether diseases are genetically determined? Lancet 1993;341:1008–9.

Epidem Rev 2002;24:137–153


