invited review

Aetiology of autism: findings and questions∗

M. Rutter

SGDP Centre, Institute of Psychiatry, Denmark Hill, London, UK

Abstract

Background Although there is good evidence that autism is a multifactorial disorder, an adequate understanding of the genetic and non-genetic causes has yet to be achieved.

Methods Empirical research findings and conceptual reviews are reviewed with respect to evidence on possible causal influences.

Results Much the strongest evidence concerns the importance of susceptibility genes, but such genes have yet to be identified. Specific somatic conditions (such as tuberous sclerosis and the fragile X anomaly) account for a small proportion of cases. Over recent decades there has been a major rise in the rate of diagnosed autism. The main explanation for this rise is to be found in better ascertainment and a broadening of the diagnostic concept. Nevertheless, some degree of true rise cannot be firmly excluded. However, the epidemiological evidence on the main hypothesized environmental explanation, namely the measles-mumps-rubella vaccine, is consistently negative.

Conclusion Progress on the elucidation of the causes of autism will be crucially dependent on the combination of epidemiology with more basic science laboratory studies.

Keywords autism spectrum disorders, epidemiology, genetic influences, incidence, measles-mumps-rubella vaccine

Introduction

Before turning to what is known on the aetiology of autism, it is necessary to note that, over the last few decades, there have been important changes in the diagnostic concept. During the 1960s, it was generally assumed that autism was a rare, seriously handicapping disorder, usually associated with intellectual disability (ID), constituting a condition that was qualitatively distinct from variations in social and communicative competence within the normal range. Both epidemiological and genetic research findings have forced a change in concept as a result of the evidence that autistic-like abnormalities can and do occur in individuals of normal intelligence (see Rutter 2004). There appears to be a broad spectrum of disorders that are closely similar in quality but milder in some respects and often occurring in individuals of normal intelligence. Even further outside the core is a group of much milder, but apparently similar, conditions that have come to be termed the ‘broader phenotype’. They are found in some one in five first-degree relatives of individuals with autism.

But outside even this broad spectrum there are other disorders that may sometimes be confused with autism. First, there is the well-defined condition called Rett syndrome after the physician who first described it (Rett 1966; Hagberg et al. 1983); it has...
been shown to be a result, in almost all cases, of a mutant gene on the X chromosome. Its downward clinical course and neurological features mark it out as different from autism but many of those who suffer from this condition go through a phase that involves social impairments superficially similar to autism. In addition, there are several less well-defined quasi-autistic patterns, similar to autism in many respects, but different in some features. These include the patterns seen in some congenitally blind children (Hobson et al. 1999) and some children reared in seriously depriving institutions (Rutter et al. 1999); in some young adults who had a serious developmental disorder of receptive language when young (Clegg et al. 2005; Howlin et al. 2000), and in some children with a semantic pragmatic language disorder (Bishop et al. 1995; Bishop 2002). In the remainder of this paper I will confine attention to what is known about the aetiology of autism and the more closely associated autism spectrum disorders (ASD), and for the most part will not deal with either the broader phenotype or these varied more atypical patterns.

**Susceptibility genes for autism spectrum disorders**

Much the best-established risk factor for ASD is genetic liability (Rutter 2005). In idiopathic cases without a known medical cause, twin studies have shown a concordance rate in monozygotic (identical) twin pairs of about 60% as compared with a rate of 5% in dizygotic (fraternal) pairs. Taken together with the population base rate for autism, this implies that the heritability or underlying genetic liability is about 90% – the highest figure among all multifactorial child psychiatric disorders. However, the twin data also show that the genetic liability extends well beyond the traditional core diagnosis of autism to include a wider spectrum of autistic-like disorders including the broader phenotype. In other words, in twin pairs discordant for core autism, the monozygotic concordance for the broader phenotype far exceeds the dizygotic concordance.

Family studies tell a similar story. The rate of ASD in the siblings of individuals with autism is about 6% – a rate many times higher than the rate of about 0.5% in the general population. The rate of 6% seems low in absolute terms but, relative to the general population, it is very high. Putting together the twin and family findings, the figures indicate that it is likely that there are somewhere between 3 and 12 susceptibility genes for autism that act synergistically. The reason why most siblings do not have ASD, despite the high heritability, is that they have only some of the relevant genes.

**Susceptibility genes for autism**

A major growth area in psychiatric genetics as applied to ASD concerns the search, using linkage and association strategies, to identify susceptibility genes (Rutter 2005). There are very promising, partially replicated, findings with respect to loci on chromosomes 2 and 7, as well as leads on other chromosome locations. It may be anticipated, with some confidence, that the actual susceptibility genes will be determined during the next decade, if not rather earlier than that. Linkage strategies study the degree to which affected members in the same family show coinheritance of the same genetic loci on particular chromosomes. Association strategies, by contrast, determine whether individuals with ASD differ from controls in their pattern of allelic variations of specific genes. That is, each person inherits one out of several possible allelic copies of each gene. Some copies will carry risk whereas others will not. Note that the susceptibility genes may well turn out to be common genes that, on their own, do not directly cause disease – rather than rare pathogenic mutations. Although it may be expected that the identification of susceptibility genes will be enormously helpful in shaping the biological research that will determine the neural basis of autism, it is much less likely that the genes themselves will be of much practical utility in terms of either screening or diagnosis.

**Single gene conditions**

The only single gene condition with an established association with ASD is tuberous sclerosis (Smalley 1998). The best estimates suggest that this is found in about 1% to 3% of cases of ASD (Harrison & Bolton 1997). Although the association with ASD is well established, its meaning in terms of the causal mechanisms that are involved is less well understood. It is probably important that the association with
ASD is particularly marked only when tuberous sclerosis is associated with severe ID, severe epilepsy and the location of tubers in the temporal lobe (Bolton et al. 2002). In other words, the implication is that the causal pathway may involve the type and location of the pathophysiology of brain disturbance deriving from tuberous sclerosis, rather than from anything directly associated with the genes on chromosome 9 and chromosome 16 that give rise to tuberous sclerosis.

Chromosome anomalies

Most of the evidence on connections between chromosome anomalies and ASD come from isolated case reports (Gillberg 1998). These are of little use in testing causal hypotheses and greater reliance needs to be placed on systematic studies of either general population or clinic samples. Initially, the strongest claims concerned the supposed association between the Fragile X anomaly and autism (Gillberg & Wahlström 1985). The initial claims of a strong association were based on unsatisfactory cell culture methods and, once DNA methods became available, it was evident that the cytological identification of fragile sites led to many false positives (Gurling et al. 1997). Systematic surveys of large samples of individuals with an ASD have shown that only about 2% to 3% show the Fragile X anomaly (Bailey et al. 1993; Chakrabarti & Fombonne 2001). This is still a meaningful and significant association but it is evident that it accounts for a very small proportion of cases of ASD. On the other hand, surveys of individuals known to have the Fragile X anomaly have shown that quite a high proportion show social and communicative abnormalities of a kind that could be confused with autism, even though they do not meet the usual criteria for an ASD (Reiss & Dant 2003). The only other chromosome anomaly at all commonly associated with ASD concerns the maternally transmitted interstitial duplications of chromosome 15 (Folstein & Rosen-Sheidley 2001). Systematic surveys of chromosome anomalies in a series of individuals with an ASD have shown that approximately 5% show anomalies of one kind or another. These are quite varied and, in most cases, their clinical significance remains uncertain.

Medical conditions and autism

There have been various debates in the literature concerning the frequency with which ASD are associated with definite diagnosable medical conditions that are likely to have been implicated in the causal processes (Rutter et al. 1994; Gillberg & Coleman 1996). One of the major difficulties in coming to a specific figure concerns the major extent to which findings are likely to be influenced by both the nature of the samples investigated and the thoroughness of the medical investigations undertaken. However, a reasonable estimate would be that something in the order of 10% of individuals with ASD has some potentially relevant identifiable somatic disease or disorder. This means that an appropriately thorough medical assessment is essential in all cases. The general consensus would be that this should include careful medical examination, including the use of Wood’s light, in order to detect tuberous sclerosis, that karyotyping should be routinely undertaken, but also that this should include the use of DNA methods to diagnose the Fragile X anomaly. Although some Scandinavian researchers have advocated intrusive further investigations as a routine – including lumbar puncture, brain scanning using a general anaesthetic, etc. (Gillberg 1990), the evidence would seem to suggest a more conservative approach in which the extent, and type, of medical investigations are determined on the basis of the clinical history and clinical examination findings in the individual patient.

Prenatal influences

Intra-uterine infections and toxins

Although the main research attention has focused on genetic influences in autism and on associated medical conditions, the evidence is clear cut that most ASD constitute multifactorial disorders. That means that some kinds of non-genetic factors are also likely to play a part in aetiology, even if we know little about them.

Isolated case reports have suggested that a range of possible intrauterine infections and toxins could play a contributory causal role in the development of ASD in individual cases (Nelson 1991; Rodier & Hyman 1998; Folstein & Rosen-Sheidley 2001; Medical Research Council 2001). These include various possible maternal circumstances that could affect the
foetus including hypothyroidism, thalidomide use, valproic acid use, cocaine or alcohol use, and congenital cytomegalovirus infection. None of these have been prominent in any of the epidemiological studies of ASD and it seems unlikely that they constitute commonly operating risk factors for ASD.

The only other established link is that between congenital rubella and autism (Chess et al. 1971; Chess 1977). Findings from a systematically studied large sample of children with congenital rubella showed that a substantial minority developed some form of ASD. The rate was substantially higher in the children whose handicaps included marked ID as well as visual and hearing defects (but ASD was not confined to that group). It is noteworthy, however, that the follow-up showed that the course of ASD in these children tended, on the whole, to be rather different from that associated with idiopathic autism in that, although the children remained markedly handicapped at they grew older, the autistic features tended to diminish. The findings, of course, are of very limited contemporary relevance in view of the rarity of congenital rubella following the establishment of population-wide vaccination programmes.

Obstetric complications and birth order

Since the first pulling together of the evidence (Deykin & MacMahon 1980), it has generally been found that autism tends to be more common in the firstborn in sibships of two but more common in the last born in larger sibships. Also it has usually been found that obstetric complications (generally of a mild variety) are more common in individuals with autism than in their unaffected siblings or in controls (Bolton et al. 1997). The balance of the evidence suggests that the obstetric complications do not constitute an environmentally mediated risk; rather, they may reflect a response to a genetically abnormal foetus (Bolton et al. 1997).

Parental social class and country of origin

Following Kanner’s initial report that children with autism were disproportionately likely to have parents of a high socio-economic background (Kanner 1943), and Lotter’s finding that there was a slight tendency of this kind (Lotter 1966), there was an interest in the possibility that high social class was associated with autism. However, both Wing (1980) and Schopler et al. (1979) did not find this. Their studies were both flawed by a failure to take ethnicity into account. Most recent surveys have not examined social background systematically but, with one possible exception (Fombonne et al. 2001), the few that have done so have not found any association. Most reviewers have therefore concluded that probably there is no association other than that associated with referral bias (MRC 2001; Fombonne 2003). That is likely to be true but it has to be said that the evidential base for the conclusion is weak. In addition, there has been some indication that autism may be more common in the UK in children born to parents of Afro-Caribbean background (Wing 1980; Goodman & Richards 1995) and Gillberg and Gillberg reported an increase in Sweden for children born to immigrant parents (Gillberg & Gillberg 1996). As with social class, the findings are contradictory, inconclusive and based on small numbers (MRC 2001; Fombonne 2003). The conclusion has to be ‘not proven’.

Monozygotic twinning as a risk factor

Non-geneticists tend to assume that the non-genetic factors involved in the aetiology of ASD must necessarily involve some form of specific environmental risk. It is important to appreciate that that is not necessarily the case (Molenaar et al. 1993; Jensen 1997). For example, Greenberg et al. (2001), and also Betancur et al. (2002), reported an apparent excess of twins among affected sibling pairs with autism (but see Hodge et al. 2002; Vischer 2002). If this finding were to prove valid, it would suggest that being a twin constituted a risk factor for autism. That could come about either because twinning is associated with an increased risk of obstetric complications or because monozygotic twinning itself constitutes a form of congenital anomaly (Hall 2003). Congenital anomalies have been found to be more common in individuals with autism and these probably index the ways in which development, which is probabilistic rather than deterministically programmed, may go awry (Vogel & Motulsky 1997). Thus, congenital anomalies are more common in twins than in singletons and are more common in children born to older mothers than in those born to younger ones (Myrianthopoulos & Melnick 1977; Rutter et al. 1990). Accordingly, it could be that these semi-random developmental per-
urbations could enhance the adverse effects of a genetic liability to ASD. It should be noted, nevertheless, that congenital anomalies show an increased rate in a wide range of psychiatric disorders, so that the risk is by no means specific to ASD.

More importantly, some scepticism is necessary with respect to the supposed finding that the rate of twinning is actually increased in ASD. Ascertainment biases are likely to have played a major role and it is noteworthy that the most systematic twin sample of Bailey et al. (1995) did not include a significant excess of monozygotic twins; nor did Hallmayer et al.'s 2002 Australian twin sample. It may be concluded that the postulated increased risk for autism associated with being a twin remains a speculative suggestion and, on balance, the evidence indicates that it is not likely that being a monozygotic twin constitutes a major risk factor.

Postnatal risk influences

All the evidence suggests that it is rare for postnatal somatic disease to give rise to an ASD. There are isolated case reports of herpes encephalitis causing autism (Gillberg 1986; Ghaziuddin et al. 1992) but this is decidedly unusual and ASD has not been reported as a common consequence of encephalitis in childhood (Rantala et al. 1991).

Measles-mumps-rubella and thimerosal

During the last decade, the main focus within the realm of possible postnatal risk factors for ASD has been on the possibility that immunization constitutes a contributory factor for ASD. First, there was the suggestion that the measles-mumps-rubella (MMR) vaccine was responsible for the huge recent rise in the rate of diagnosed ASD (Rutter 2004). It was argued that through the route of a vaccine-caused gut disorder, there was leakage of protein products into the blood stream and that these then caused a special regressive form of autism (in which there was a loss of previously acquired social and communicative skills). A range of epidemiological studies was undertaken to determine whether the use of the MMR vaccine might be responsible for the worldwide rise in the rate of autism as diagnosed, and in particular whether it led to this postulated regressive form of autism. The evidence is consistently against the MMR hypothesis. If MMR had been responsible for the rise in ASD it would be expected that the introduction of the vaccine, in countries in which the take-up was rapid and very high (as was the case in the UK), should be followed by a large step-wise increase in ASD, that this should be followed by a plateau in rate, and that when MMR was stopped (as it was in Japan) this should be followed by a fall in rate. In all these phases, the main changes should apply to regressive autism. The evidence shows that none of these expectations were borne out (Rutter 2004). The hypothesis regarding Thimerosal (a preservative that was, until recently, used in many vaccines) is somewhat different in detail, in that mercury is known to be a neurotoxin; accordingly, a direct adverse effect on the brain was expected. However, it remains uncertain whether a ‘bolus’ effect causes damage (i.e. the immediate, large, but very transient rise in mercury level following vaccination) or whether the damage derives from the cumulative mercury build-up resulting from multiple vaccinations. The epidemiological evidence on Thimerosal is much less than that on MMR but again the findings are negative.

These negative conclusions give rise to two main queries. First, if neither MMR nor Thimerosal is responsible for the rise in autism, what has caused the increase? It is clear that the main explanation is that it derives from a combination of better ascertainment and a broadening of the diagnostic concept. However, the possibility that, in addition, there has been a true rise in incidence because of some, as yet unidentified, environmental risk factor cannot be ruled out (Rutter 2004). Second, although it is no longer plausible that MMR or Thimerosal have led to an overall increase in ASD, the epidemiological data cannot exclude the possibility that either might have a risk effect in a small proportion of unusually susceptible children. There is no evidence supporting this suggestion but it cannot be firmly excluded.

Other possibilities

It needs to be added that there are other possible causes of ASD. For example, there are uncertain pointers to the role of immunological abnormalities. The supporting evidence is weak but it is a group of risk factors that warrant further exploration.
Conclusions

In summary, there are good epidemiological data indicating that the true incidence of ASD now is likely to be of the order of 30–60 cases per 10,000, as compared with the original estimate of four per 10,000 made some four decades ago (Rutter 2004). Administrative data show massive increases over time in the rate of diagnosed ASD and it is clear that, in large part, this is because of the combination of better ascertainment and a broadening of the diagnostic concept, but a true rise over time in the incidence of ASD cannot be entirely ruled out. Despite strong claims made about the possible role of MMR in relation to the causation of autism, there is no convincing evidence in support of this hypothesis. In particular, the rate of ASD shows no particular association with either the stopping or starting of MMR and there has been no change over time in the pattern of association between ASD and either bowel disturbance or developmental regression. The evidence with respect to a possible association with Thimerosal, a preservative in some vaccines, is much more limited but, again, there is no supporting epidemiological evidence of a causal association. It remains possible that there has been a true rise in incidence because of some environmental risk factor but, if so, it remains quite obscure as to what that factor might be.

The genetic evidence is clear cut that ASD are multifactorial conditions caused by multiple genes and some, as yet to be identified, non-genetic factors. The genetic factors that underlie ASD are likely to be heterogeneous but it remains unclear whether that heterogeneity is indexed by clinical features and, if it is, which they are. The evidence is also clear-cut that the genetic liability to ASD involves a broader phenotype that extends well beyond the traditional diagnosis of a handicapping condition of autism. However, it is significant that the broader phenotype does not seem to be associated with either epilepsy or ID and very little is known on the factors, genetic or non-genetic, that are implicated in the transition from the milder broader phenotype to a seriously handicapping disorder.

Epidemiological findings have been helpful in both ruling in and ruling out various postulated causal influences and they will continue to be formative in that connection. Nevertheless, it is evident that progress is going to be crucially dependent on the combination and integration of epidemiology with more basic science laboratory studies.

References


*Accepted 14 July 2004*