The Inheritance of MS Susceptibility

D. Lord[1], B.Sc., A. G. O'Farrell,[1] M.Sc., Ph.D., M.R.I.A.,
H. Staunton[2], B.Sc., F.R.C.P., F.R.C.P.I., Ph.D.,
and E. Keelan[2], M.B., M.R.C.P.I.

[1]Maynooth College, Co. Kildare, Ireland.
[2]Richmond Institute of Neurology and Neurosurgery, Beaumont Hospital, Dublin 7, Ireland.

Summary

The object of this study was to determine, if possible, the mode of inheritance of the susceptibility to multiple sclerosis (MS).

It was known that no single-gene model could fit the observations, so oligogenic models (models involving a small number of genes) were sought.

Oligogenic hypotheses were tested against the available population data for MS until a reasonable level of agreement was found. The best-fitting simple hypothesis was this : MS occurs only in people who are homozygous for a recessive gene and carry a dominant X gene, and then only with reduced penetrance.

The dangerous allele m^- at the autosomal locus appears to be fairly uniformly distributed across England, Ireland and Canada, occurring in 10-30% of the gene pool. There are large variations in the frequency of the allele s^- at the X-locus, ranging from 10% up to 72% of the gene pool. The penetrance varies significantly with geographical location, but nowhere approaches unity, so that the environmental factors are of great importance.

The hypothesis explains the broad features of the population pattern of the occurrence of MS and it gives an outstanding fit to the best available data on MS in relatives.

The result may assist attempts to map the genetic data on MS, opens the way for a reassessment of the attempts to identify the environmental factors, and it makes possible the completion of non-empirical risk tables for various countries.

Similar techniques may be applied to other disorders with a genetic component in their aetiology

1. Introduction

While the aetiology of multiple sclerosis (MS) is unknown, the presence of an inherited susceptibility is supported by the observation that first degree relatives of patients have substantially increased risk, the prevalence being higher the closer the relationship.[1] Although this is further supported by comparison between monozygotic (MZ) and dizygotic (DZ) twins,[2] a MZ twin concordance rate significantly less than 100% indicates a major environmental factor. This was borne out by studies of MS in migrants,[3-5] which supported the hypothesis that events in childhood or adolescence are implicated.

The putative genetic factor in MS has certain features of autosomal recessive inheritance.[6] However, no simple genetic hypothesis fits the observed population pattern.[7-11,2] Thus, a polygenic mechanism may operate, which may also include a sex factor. Further, there is pedigree evidence to suggest a locus mapping near the HLA (human leukocyte antigen) system on chromosomal 6, and, in addition, the Ig (immunoglobulin) region of chromosome 14.[10] However, despite widespread interest and considerable detailed study, no satisfactory hypothesis has emerged to explain the population pattern.

In such situations where a genetic factor appears definite, but not the major factor and not easily definable, some formal mathematical treatment is required in its analysis. The aim of this study was to use such a method to test various oligogenic hypotheses against the data from published population studies in order to find the "best fit".

The study was concerned only with the genetic aspect of the disease, and no attempt was made to investigate the environmental aspect. This is not to say that it was ignored. Its existence was allowed for, but its character was not analysed.

2. Material

(2.1) *General analysis of data*

No new experimental data are presented in this study which is rather an analysis of already existing data. Before describing these, some preliminary remarks are in order.

First, it is necessary to make precise what we mean by MS, in view of the extraordinary diversity shown by the course of the disease, and the consequent problems in ascertainment which bedevil all population studies. The normal criteria for the diagnosis of MS are Schumacher's,[12] or some variation on them. In these, the most important are the requirements that there be objective evidence of white-matter lesions separated in space, and that there be evidence of remission and relapse or slow and stepwise progression. However, the diagnosis of MS is sometimes made at autopsy, in otherwise asymptomatic (or "sub-clinically symptomatic") cases, on the basis of the occurrence of disseminated sclerotic plaques. Also cases are sometimes suggested as possible MS on the basis of characteristic MRI scans, in the absence of other symptoms.[2] Further, the distinction is made, under Schumacher's criteria, between three categories of clinical MS: definite, probable, and possible. Finally, it is by no means certain that the clinical conditions classified as MS are all due to a single homogenous disorder. In particular, there is probably confusion of MS with some types of optic neuritis.

The various field studies referred to below adopted varying approaches to this question. Most workers in recent times have used Schumacher's criteria. Sometimes the definite, probable and possible categories were distinguished, and sometimes they were grouped. Some studies included cases that did not meet the Schumacher criteria. In most cases,

the diagnoses were made or checked by neurologists, but not in all. These variations, and the variations in practice, and in experience of MS, between countries, make for difficulties in comparing data.

Looking for an invariant and objective definition of the term "MS case", we considered adopting one of the following:

(MS1): An individual is an MS case if, sometime in life, he or she has been or will be diagnosed by a neurologist as having clinically definite MS, according to Schumacher's criteria.

(MS2): An individual is an MS case if, sometime in life, he or she has been or will be diagnosed by a neurologist as having clinically definite or probable MS, according to Schumacher's criteria.

(MS3): An individual is an MS case if, sometime in life, he or she has been or will be diagnosed by a neurologist as having clinically definite or probable or possible MS according to Schumacher's criteria.

(MS4): An individual is an MS case if he has or will satisfy Schumacher's criteria whether or not he is seen by a neurologist.

(MS5): An individual is an MS case if he has or will satisfy Schumacher's criteria, or shows or will show characteristic MRI scan, or shows or will show disseminated lesions at autopsy.

Depending on the definition adopted, there will be more or fewer MS cases in each population. This would be more of a problem if MS were a purely genetic disease. Since, in fact, any genetic factors certainly act with reduced penetrance (see next section), the effect of selecting one or another definition will simply be to alter the penetrance. In fact, the proportion of cases that reach clinics possibly does vary from country to country, and the proportion of cases classified as possible certainly shows bias from study to study (see Appendix 1). However, and this is a crucial point for our analysis, there is no evidence that the proportion of cases classified in one or another way varies from subgroup to subgroup *within a given study*. In particular, the allocation of men and women to the categories probable and possible is consistent (see Appendix 1).

Definitions like MS4 and MS5 make the concept of "MS case" impervious to measurement, so choice from MS1-3 was indicated.

We adopted definition MS3 as the definition of an MS case. This has two practical advantages over MS1 and MS2. First, there are more cases in small samples, and this helps to smooth random irregularities. More importantly, survival data are important, as will appear below, and even in MS2 cases these are normally given in years from age of onset rather than from age of clinically definite/probable MS diagnosis, or even from age of presentation. At onset, an MS2 diagnosis is impossible, since recurrence in time is a prerequisite.

This definition was designed to get away from the (at present) ill-defined question of whether this or that individual "really" has (or had) MS. As long as we do not understand the aetiology, we are in no position to say whether or not a person "really" has MS. Note that the definition excludes as MS cases all those who die before diagnosis, whether or not there is post-mortem evidence.

To test genetic hypotheses, the primary data needed are the *frequencies* of MS cases in various groups, such as Irish men, British women, sisters of Canadian patients, etc.

The **frequency** of a condition in a group is defined to be the probability that a person born into that group will develop the condition. The frequency of MS3 in a group is the proportion of members of that group who will ultimately prove to be MS3 cases. (It is impossible to know that a living subject is not an MS3 case. To estimate the proportion of MS3 cases in a group, one must make allowance for the residual risk to living members of the group.)

In practice, most studies report *prevalence* or *incidence* rates, and these have to be converted to estimates of frequencies.

The **prevalence,** or point prevalence, of a condition in a group at an instant in time is the proportion of the group who are known to have the condition at that time. In the case of MS, the prevalence of MS3, in, say Northern Ireland males, on the First of January, 1989, is the number of live males on the provincial MS register on that date who are classified as having definite, probable or possible MS, divided by the number of males alive in the province on that day. Prevalence studies are normally carried out by examining registers or searching clinical files. In practice, prevalences are frequently estimated statistically, by sampling, rather than directly measured. In large populations, this may even be a *preferable* way to estimate the true situation, in view of the tendency for underestimation in large-scale studies (see remarks in Bauer 1987). Prevalences are often reported in numbers per head of population in the age-range 20-60.

The **incidence** (per annum) or attack rate of a condition in a group is the proportion of the group who manifest the condition for the first time in the course of a given year. In the case of MS, the incidence of MS3 in a given group and year is the per capita number of new cases classified in that year. Incidence estimates are usually based on studies of consecutive admissions to clinics, or on annual registrations in national registers.

It is theoretically possible to derive frequency data either from prevalence data or from incidence data, provided these are comprehensive, and provided detailed birth-rate, age-structure and survival data are available for the group in question. In practice, such data are rarely available, and in any case, the estimates of prevalence or incidence are subject to substantial error, so that the use of the detailed formulas is either impossible or inappropriate. Thus it is normal to use the rules of thumb[13]

$$\text{mean incidence} \sim \text{mean prevalence} / \text{mean duration},$$

$$\text{mean frequency} \sim \text{mean incidence} / \text{mean birthrate}.$$

Here the mean duration refers to the survival in years after diagnosis, and the mean birthrate refers to the birthrate per capita per annum in the group. These approximate formulas become exact in a steady-state situation, with constant incidence and birthrate.

Early population studies on MS reported data such as the proportion of death certificates giving MS as a cause and the proportion of hospital admis-

sions in which MS is the diagnosed ailment. These are difficult to convert into frequencies with any degree of reliability. One is obliged to use methods such as that of Kurtzke,[14] who took the ratio of admissions for MS to admissions for another disease (amyotrophic lateral sclerosis) that is assumed to have an invariant incidence worldwide.

A few studies have reported direct frequency estimates, not derived from incidences or penetrances. Usually these were based on analysis of registers of MS patients or other categories (such as twins) by birth date.

Many studies have failed to report separately on men and women, and others have referred to relatives, without specifying the kind or degree of relationship, or even the number of relatives studied. These are virtually useless for our purpose. Because of the comparative rarity of the disease, large scale studies are needed to produce reliable data. All available family studies involved relatives of fewer than 1000 patients and so allow only a limited degree of discrimination between oligogenic hypotheses. On the other hand, it was a key point in our procedure not to ignore any reasonably reliable datum, even if it consisted of a single item, such as the relative risk to men and women in some country. Standard pedigree analysis suffers from a weakness in dealing with rare oligogenic conditions, in that the cases studied are usually drawn from a single geographical area, and the effects of some genes may be undetectable.

2.2 *Data Sources*

1. A number of important sources provide systemic reviews and compilations of available data on MS.[15,14,10,9,1,16,11,17]

The studies listed below (country of origin nominated), provide additional detailed raw data on which computations were performed.

2. Iceland[18]
3. Northern Ireland[19]
4. Northern Ireland[20]
5. England[13]
6. Scotland[21]
7. Orkney[7,22]
8. Canada[23,24]
9. Canada[25]
10. Australia[4]
11. England[8]
12. Faroes[26]
13. U.S.A.[27]
14. Canada[2]
15. South Africa[3]
16. Ireland[28]

The Northern Irish data, the English data from Schapira et al, and the Canadian studies were the most important for our purposes.

Some specific data derived from the foregoing studies are as follows:

2.1.0 From Matthews et al.:[1]

Estimated frequencies for MS range from a high of 1/215 in Orkney to practically zero in some equatorial locations. Typical values in large high-risk areas are about 0.1%.

More recent MS studies, and especially re-surveys of an area previously surveyed, tend to produce higher prevalence estimates (p.5).

Case-finding is more complete in smaller samples (p.6).

Mean age at onset and the course of the disease are similar whether the disease is common or rare in a locality (p.7). Onset is, on average, approximately four years later in men than in women. The mean age at onset in women is about 30 (p.9, Table 1.5). Also, the age of peak incidence is lower in women (p.7).

The siblings of MS patients are more likely to develop MS than either the parents or the children of MS patients (p.74).

The disease is more prevalent in women than men. Furthermore, the estimates obtained for the ratio of female : male incidence vary considerably from population to population (Table 4, p.7) from a low of 1.1 : 1 in Iceland, through 1.2 : 1 in Northern Ireland, 1.8 : 1 in the USA as a whole, 2.0 : 1 in New Orleans, 2.5 : 1 in South Africa (native whites), to 2.8 : 1 in Perth, Western Australia. The best available estimate of the global F : M prevalence ratio in populations in which MS occurrence is significant is 1.4. The F : M prevalence ratio in a region is usually different from the F: M incidence ratio, due to different ages of onset and survival times.

These observations are supported by further data reported in other sources, especially Bauer,[17] and by more specific data quoted below.

2.1.1 From Spielman and Nathanson:[10]

They quote data from some older studies (p. 47, Table 1), on the occurrence of MS in sibs and parents of MS cases. These include studies

(1) from Germany, in 1937, which found MS in 4 of 444 sibs and 0 of 212 parents;

(2) from England, in 1951, which found MS in 6 of 609 sibs and 7 of 1448 parents; and

(3) from Sweden, in 1953, which found MS in 22 of 2815 sibs and 5 of 1493 parents.

Spielman and Nathanson describe a number of pedigree studies, most of which focused on the HLA associations and linkage of MS. The most significant result was that (p.54) of

(4) Stewart et al. who obtained a result in 1981 which provides strong evidence for linkage to the HLA-DR locus. However, the data were consistent either with dominant or recessive action of an MS susceptibility gene very near the HLA-DR locus. In the dominant case, the maximum likelihood estimate of the gene frequency is 0.2, whereas in the recessive case, the frequency is 0.6. Stewart et al. favour the dominant action for the gene near HLA-DR since it is supported by the distribution of HLA-DR2 genotypes in MS cases and by pedigree analysis incorporating HLA. The most persuasive argument for dominant action is based on the analysis of homo- and hetero-zygosity for the allele HLA-DR2 in 60 patients with definite MS. Knowing the frequency of DR2 in MS patients, and assuming tight linkage (recombination fraction $r=0$), predicted frequencies for the DR2 genotypes were computed and compared with observation:

HLR-DR alleles	actual	Expected dominant	Expected recessive
2/2	3	4.9	7.7
2/other	37	35.1	27.6
other/other	20	20	24.7
χ^2, 1 d.f.:		0.85	6.96
Probability:		>0.3	<0.01

The risk to DR2 carriers is 2.5-4 times that of the general population, but at most 50-60% of MS cases carry DR2. Different DR associations are found in some parts of the world.

They also quoted

(5) Pandey et al, who in a 1981 study of 55 patients found indications of a recessively-acting allele at a locus near GM on the immunoglobulin G heavy chain on chromosome 14.

2.1.2 Iceland (Gudmundsson):[18] A re-evaluation by K. R. Gudmundsson of a survey which was originally carried out in 1962 by Gudmundsson and Gudmundsson, covering the period 1946-55. Schumacher's criteria were used. The number of patients with multiple sclerosis who were alive at some time during the survey period 1946-55 was 104 . (This compared with a measurement of 80 cases in the earlier survey). Of these, 90 were classified as having clinically definite MS, 4 as probable, and 10 as possible. Of the total material 41 patients were males and 63 females. The female : male ratio was therefore 1.54. The deviation of this ratio from 1 was significant at the 95% level. The MS3 prevalence per 100000 of population was 57.1 (F68.2 M46.1), the MS2 prevalence was 52.0 (F65.7 M38.6).

Direct measurement of incidence gave $3.32/10^5$ in females, and $3.35/10^5$ in males, taking all cases; 3.13 in females and 2.93 in males, taking only definite and probable cases. The rate was highest in the age-group 20-40, without significant difference between the sexes.

The re-evaluation of the 1962 survey involved the disqualification of one definite case and one possible case, besides the addition of 26 cases.

The prevalence rates were calculated on a population (1955) of 89,325 males and 79,155 females. The incidence rates were calculated on a population (1950) between the ages of 10 and 59 of 48,446 males and 47,162 females.

Detailed figures on the durations were provided (cf. Appendix 2).

The familial cases among the 104 cases were as follows:

father-daughter
2x brother-sister
sister-sister
uncle-nephew
aunt-niece
first cousins F+F
first cousins F+M

He summarised (p.14) some earlier studies of the variations in prevalence by sex. Excluding those listed below, these gave the following values for the percentage females among cases:

study	% female
Sweden 1949 Müller:	59.9
Great Britain 1952 McAlpine and Compston:	64.2
England 1963 Poskanzer et al.:	58.9
Denmark 1956 Kurland and Westlund:	56.6
Scotland 1950 Adams et al.:	55.6
Canada 1964 Stazio et al.:	69.8

2.1.3 Northern Ireland (Ashitey and Millar 1970):[19] This study produced direct estimates of frequency of MS in those born in the years 1901-25 in the various regions of Northern Ireland. The study involved the whole population (1,480,000). Pre-Shumacher criteria (Allison-Millar criteria 1954) were used,[29] involving categories probable, possible and early. The majority of the cases were diagnosed by the authors, so a high degree of uniformity in the diagnoses is likely. On the First of October 1968 there were 1418 multiple sclerosis cases, 1238 of whom were born in N. Ireland. A detailed study was made of 783 cases born in N. Ireland from 1901 to 1925. It was concluded that the risk of developing MS is not independent of place of birth (the risk was significantly lower for those born in Belfast), but is independent of the year of birth of the patients. The average frequency of MS was 1.06 per 1000 live births. The frequency in women was 1.21/1000 and in men 0.91/1000, for a ratio of 1.33. The raw numbers were 448 female cases and 335 male, out of about 738,700 births.

2.1.4 Northern Ireland (Millar and Allison):[20] This study attempted to trace all cases of MS in Northern Ireland, and identify all familial cases. Allison-Millar criteria were used. The majority of cases were seen by the authors, but some diagnoses were based on second-hand accounts. Seven hundred cases were found, and in 44 out of 668 families more than one case was found. The familial cases were as follows: (each line lists a family type, preceded by the number of the type found):

1 × grandfather, father and son
5 × father and son
1 × mother and daughter
1 × mother and 2 daughters
1 × mother and son
1 × mother, her son and daughter
13 × brother and sister
1 × brother and sister and the sister's daughter
3 × 2 brothers and their sister
1 × 3 brothers
8 × 2 sisters
1 × 2 sisters and their brother
2 × uncle and nephew
2 × aunt and niece
3 × first cousins

They provided case histories and family trees of all cases.

They gave data on family size (Table 4, p.36), as follows: 36 families had 1 child, 60 had 2, 70 had 3, 100 had 4, 86 had 5, 97 had 6, 75 had 7, 64 had 8, 42 had 9, 21 had 10, 6 had 11, 6 had 12, 4 had 13, and 1 had 14.

There were 307 males and 395 females with MS out of 1807 males and 1800 females in their families. The average family size was 3607/668=5.40

They observe that no simple genetic mechanism explains the pattern they found, and continue (p.38): "It is always tempting to proceed logically at this stage and to say that perhaps two genes are involved, and by juggling with possible combinations of one or two dominant and/or recessive genes to demonstrate that one's own observations fit in with some theoretical hypothesis. For many reasons this is a dangerous exercise".

2.1.5 United Kingdom (Shapira et al.):[13] This study produced estimates of frequencies of MS in siblings and parents of MS patients. The study was part of an investigation of 700 consecutive patients referred by general practitioners in Northumberland and Durham as cases of multiple sclerosis. Of these 65 were rejected because the diagnosis was considered incorrect or because they lived outside the two counties named. Information was obtained on the parents

and siblings of the remaining 635 patients. Criteria approximating Schumacher's were used, with categories probable (evidence of dissemination of lesions in time and space), latent and possible (p.319). The unbiased method of "multiple ascertainment" was used in computing frequencies (i.e. each family was counted as many times as it occurred in the original list of propositi). The reported estimates (Table 1) are divided by sex, a particularly important feature for our purposes. (We have corrected an arithmetical slip in the computation of the risk to sisters of female cases, as presented in Table IV, p.322). The estimates are not age-adjusted.

TABLE I: North of England

frequency of MS in:	/10^5
women	136±55
men	105±42
sisters of female cases	1964±561
brothers of female cases	846±377
sisters of male cases	1282±520
brothers of male cases	416±293
mothers of cases	829±369
fathers of cases	332±234

From data in Shapira et al. 1963[13]

Some information on cousins was also obtained. The information was regarded as satisfactory in the case of 229 families, yielding 2203 cousins, 35 of whom were cousins of families with 2 affected sibs. The frequency in first cousins of propositi was 27.2/10^5. It was felt that this might overestimate, since cousins with the disease were thought to be more likely to be reported.

They quoted durations of 19.5 years in women and 19.1 years in men, for a duration ratio of 1.02. Mean age at onset was given as 36.3.

Blood-group associations were sought, and three statistically-significant associations were found, including one with the Lutheran a blood-group.

2.1.6 Scotland (Shepherd and Downie):[21] This geographically-orientated study (prevalence day 1-12-1973) produced prevalence estimates broken down by regions within Scotland.

The mean duration from onset to prevalence was 15.3 years.

The mean age of cases on prevalence day was 48.6.

Age and sex-specific rates were presented in a graph (p.311), from which one can extract the information that the mean incidence (per annum) during the period 1959-73 was approximately 5.3/10^5. In fact, reading from the graph, one has the age and sex-specific numbers of cases:

age	0-19	20-29	30-39	40-49	50-59	60-69	70+
F	20	125	250	384	370	250	80
M	20	50	130	215	280	235	55

The average incidence per annum was 187.4 women and 125.6 men, for a ratio of 1.49.

In the age-group 20-60, the averages were 282.3 women and 168.8 men for a ratio of 1.67. 95% of patients were in this age group.

It was found that men show later onset, and shorter course.

2.1.7 Orkney (Roberts et al. 1979[7], 1979[22]): This study of 53 multiple sclerosis patients in Orkney aimed to gather data for a population genetics analysis. Orkney was selected because of the record reported prevalence of MS. 51 patients (45 probable, 6 possible) and 48 matched controls were studied, and 4327 relatives traced, back to a baseline date of 1775 A.D. HLA and other serological typing was carried out. Consanguinity and kinship were mesasured, and pedigree analysis undertaken.

On the basis of the consanguinity study, the patients were significantly inbred, but not significantly more so than the controls. (This was borne out by blood-typing studies). The involvement of a rare recessive gene can be ruled out. The involvement of a widely-distributed recessive gene of frequency at least 6% in the gene pool remains a possibliity. This would be at least 3 times as high as that for the most common established recessive disorders in Britain.

The number of relatives of cases did not differ significantly from those among controls, so it was considered extremely unlikely that a recently introduced single dominant or codominant gene is responsible for the disease. The involvement of such a gene, introduced 20 generations ago, and widely distributed, remains a possibility, provided it confers some survival advantage. Alternatively, such a gene, introduced in the very remote past (the founder effect), would not need to confer an advantage.

Affected relatives of cases were 2 of 94 parents and 1 of 164 sibs (making 1 in 86 first-degree relatives), 4 of 893 second-degree relatives (1 in 223), and 3 of 825 third-degree (1 in 275). In controls, the numbers found were 1 of 329 first-degree, 0 of 829 second-degree, and 1 of 801 third-degree.

None of the pedigrees suggest simple autosomal dominant, autosomal recessive, or X-linked inheritance, and these were regarded as highly unlikely. Involvement of such genes, heavily masked by environmental factors, was considered possible. Autosomal recessive or X-linked dominant were considered more likely than autosomal dominant or X-linked recessive.

From segregation information, penetrance could be estimated as 12%.

On prevalence day 1-12-74, there were 30 female cases and 21 male, for a ratio of 1.43.

The polygenic hypothesis was discussed, and favoured (p.234). For their purposes, this is the hypothesis that liability to MS is under the control of a number of genes, interacting with one another in a complex way, but essentially so that each "dangerous" allele contributes a certain amount to a quantity, and when the quantity passes some threshold, then the subject is susceptible to the disease. The broad features of the pattern are compatible with this hypothesis, including the consistent excess of affected females. The only contra-indication is that incidence is not elevated in relatives of male cases. On the polygenic hypothesis, the quantity *hereditability* may be estimated. It is the proportion of the variance in liability to develop the disease that is due to additive genetic factors. Hereditability was estimated, giving estimates from 31% to 47%, but the numbers were considered too small for confidence. Earlier estimates of hereditability ranged from 44% to 57%.

In the second paper,[22] they reported that the patients and two sets of controls were blood-typed using 23 antisera. The only significant difference between affecteds and controls was in the Lutheran system, where those affected have an excess of the Lu^a gene. However, the

gene is rare, even in patients (7.5%, as against 2.3% in controls), and there is no evidence for linkage; there is simply an association.

A slight but significant excess of O in the ABO system was noted here, as has been noted in many other British studies. There was also an excess of ccddee in the Rhesus system.

2.1.8 Canada (Sadovnick et al, Sadovnick and Baird):[24,23] This study in British Columbia obtained data on 815 index linked cases and 11,345 of their relatives, and superseded an earlier study.[30] Age-specific risks were calculated for first, second and third degree relatives of patients. There were 7541 relatives of 539 female patients and 3804 relatives of 276 male patients producing detailed estimates, with standard deviations, of the frequency of MS in parents, siblings, children, uncles, aunts and cousins of MS cases, divided by sex. Modal pedigree size was 21-25 individuals, and information was available on an average of 14 relatives per case. The diagnostic criteria were based on Schumacher and Poser et al 1984, and were stringent, including only the categories "autopsy proven", "clinically definite", "probable" and "laboratory-supported", and excluding possibles. An important feature of the method was the careful adjustment of age-of-onset. This is the only study used that originally included adjustment for age. Calculations based on these figures yielded estimates for the (lifetime) risks to some categories of relatives (Table II). The rates for fathers and mothers were calculated from the stated F:M ratio of 1.951 :1 in probands, and the stated rates for mothers of and fathers of patients of each sex.

TABLE II
British Columbia

frequency of MS in	/10^5
women	280±98
men	143±50
sisters of female cases	5650±1100
brothers of female cases	2270±710
sisters of male cases	3460±1140
brothers of male cases	4150±1280
mothers of cases	3750±1130
fathers of cases	1590±800

From data in Sandovnick et al. 1988 and Sweeney et al. 1988[23,25].

2.1.9 Canada (Sweeney et al.):[25] This study in British Columbia aimed to estimate prevalence in the province on July 1, 1982. 239,412 neurologists files were searched by one researcher, and other sources were also used. Modified Schumacher criteria were employed (cases with onset outside the ages 10-50 were admitted; for possible MS criteria of Rose et al, and of McDonald and Halliday were used; diagnoses of optic neuritis were noted). A total of 4620 non-duplicated cases of MS were identified and classified. For the prevalence-day estimates, 3632 cases were included, 2498 women and 1134 men. The prevalence estimate for definite/probable MS was 93.3/10^5 population, and including possible MS and optic neuritis, was 130.0/10^5.

The sex-specific prevalences were, for definite/probable 126.4 in women and 59.8 in men, for possible 37.9 and 15.4, for optic neuritis 14.1 and 6.8. Of all types, the prevalence was 178.4 in women, 81.9 in men, for a ratio of 2.18.

The mean age of cases was 44.7. The mean period from diagnosis to death in those cases who had died before prevalence day was 7.8 years.

The population totalled 2,783,142, of whom 1,380,000 were male and 1,400,000 female, for a female : male ratio of 1.01. Thus the sex-specific prevalences were 178.4±3.57/10^5 in women and 81.9±2.44/10^5 in men.

Of 39 neurologists dealing with MS, 2 refused to cooperate with the study.

2.1.10 Western Australia (McCall et al).[4] This survey, using Allison-Millar criteria produced separate estimates of prevalence and incidence for Perth and some rural areas of Western Australia. Perth had a population of 420,133, and 83 cases (all categories) were found (prevalence day 1-6-61), for a prevalence of 19.8/10^5. In the rural areas, the population was 261,904, and 39 cases were found, for a prevalence of 14.9/10^5. In Perth, the mean age at onset was 32.8 in women and 34.0 in men, and the mean duration from onset to prevalence day was 10.3 years in women and 10.6 in men. Different values were obtained in the country areas. The main differences between country and city were thought to be primarily due to preferential migration of the sick to the city. The mean age of women patients was 42, and of men 44. The incidence estimates were based on the 1954 population (176,815 females and 171,832 males in Perth, 110,790 females and 128,282 males in the country), and were 1.7/10^5 in women and 0.6/10^5 in males in the city, and 1.5/10^5 in women and 0.6/10^5 in men in the country.

2.1.11 England. (Roberts and Bates).[8] A study of relatives of 206 patients in North-east England which involved documentation of over 6000 relatives. Relatives were classified as first-, second- or third-degree.

73 male patients had 5 of 447 first-, 2 of 1095 second-, and 3 of 797 third-degree affected. 133 female patients had 7 of 843 first-, 5 of 2223 second-, and 0 of 1050 third-degree affected. In all, 0.93% of first-, 0.21% of second-, and 0.11% of third-degree relatives were affected.

In families of two male patients there were three and two, respectively, affected relatives, on both sides of the families. No female patient had more than one affected relative. They considered that "this suggests that the intensity of loading of the susceptibility genes is greater in male patients, i.e. the thresh-

propositus	relative	first-degree	second-degree	weighted
M	M	38.9±18.7	46.7±34.5	44.6±16.0
M	F	67.4±11.6	40.7±35.8	68.3±10.8
F	M	41.4±13.5	48.2±24.5	43.1±11.8
F	F	56.2±9.7	52.6±21.3	55.6±8.9
all	all	52.8±6.02	48.3±13.5	3 : 53.6±39.0*w*: 52.1±5.4

old of development of the disorder is somewhat more extreme in males than in females."

They derived hereditability estimates by sex of propositus and relative:

These did not suggest that hereditability is greater in first degree than in others, and hence suggested "that factors of the immediate family environment are not important".

They said (no figures given) that estimates from sibs are not elevated over those from parents, but remark that age-adjustment would produce a spread.

2.1.12 Faroes (Kurtzke and Hyllested):[26] This is one of a number of interesting studies directed at the Faroe Islands, which were a focus for the search for an infectious agent associated with MS, following suggestions of an MS epidemic in the fifties.

The study attempted to locate all cases of MS in the Islands since about 1920. 25 cases were identified among native born resident Faroese up to 1977. All but one met Schumacher criteria.The population was 41,000, and 11 of the cases were dead by July 1977. The survivors there were 8 women and 6 men, for a prevalence ratio of 1.33.

The distribution of year of onset shows strong clustering, consistent with the theory of a point-source epidemic.

The periods (in years) from onset to prevalence day among survivors were:

Females: 18, 24, 32, 32, 30, 28, 7, 21 (average 24);

Males: 18, 22, 33, 19, 33, 29 (average 25.7).

2.1.13 USA (Kurtzke et al.):[27] This study of multiple sclerosis in U.S. veterans provided some of the most reliable evidence on the prevalence gradient with latitude. We are not concerned here with the environmental factors in MS, and we note only the following data from the study: 5305 veterans with MS were matched with controls. There were 4923 white males and 182 white females in the study, and some small groups.

The relative risk to F and M was computed by comparing their respective case/control ratios (matched controls). For white females/white males the result was 1.96 in the north, 1.68 in the middle, 1.38 in the south. The trend is not statistically significant (p 1232). The average ratio was 1.79, and adjusted for geography gives 1.73.

2.1.14 Canada (Ebers et al.):[2] This major twin study was carried out by seeking twins of 5463 patients attending 10 clinics, distributed from Nova Scotia to British Columbia.

27 MZ twin pairs and 43 DZ twin pairs were found. Seven of MZ were concordant for MS, and 1 of DZ. The risk to non-twin siblings was quoted as 1.9%.

Magnetic resonance imaging showed changes suggesting subclinical disease in two more MZ twins of cases.

Earlier twin studies were reviewed in the study. Most of these were limited in scope and suffered from methodological problems, and are believed to overestimate DZ risks. The main exception is Heltberg and Holm 1982,[31] a study which linked the national Danish twin register to the national MS register, supplemented by further investigations, and which found 47 twin pairs in which clinically-definite MS occurred, of which 4 of 19 of MZ and 1 of 28 of DZ were concordant.

2.1.15 South Africa (Dean):[3] This is a study of incidence and prevalence, with particular attention to the comparison of immigrants to native-born South Africans, and to racial and regional variation. It was found that immigrants from the UK were at significantly greater risk of MS than white, English-speaking natives. The white, English-speaking population numbered 1,374,000 males and 1,389,000 females and Dean found (Census day 1960) probable MS in 46 males and 112 females, and possible and early probable MS in 9 males and 12 females, giving prevalences for probable MS of $8.0/10^5$ in females and $3.3/10^5$ in males. Adjusted for age (using the population-structure of England and Wales), these became 9.6 in females and 4.1 in males, for a prevalence ratio of 2.34. For immigrants from the UK, the ratio of prevalence was 1.6 compared with 1.4 at home.

He also measured incidence directly, and obtained values of $0.9/10^5$ in white women and $0.4/10^5$ in white men, for a ratio of 2.25. The incidence ratio for native whites was 0.5/0.2=2.5.

Survival data are not given. But mean age at onset (among English-speaking white natives) was 31, onset to diagnosis averaged 8 years, and diagnosis to prevalence day averaged 6 years.

2.1.16 Ireland (Brady et al.):[28] A prevalence study, combined with the compilation of a register for the twenty-six counties south of the border. The prevalence day was census day, 1971. Information was gathered through general practitioners, and diagnostic criteria are not stated, although nearly all patients had been examined by a neurologist or consultant physician. They found probable MS in 818 men and 1133 women, and possible MS in 97 men and 129 women, giving overall prevalences of $65.5/10^5$ for probable and $73.1/10^5$ for probable+possible.

Age and sex-specific rates were given.

Prevalence by health board area was analysed, showing a significant ($P<0.01$) deficiency of cases in the south-east area, and some other features.

3. Methods

First, the data were analysed and compared to extract estimates of frequencies, (or life-time risks) and other related quantities. Then these quantities were combined with an oligogenic calculus to select the mode of inheritance which best explained them. Finally, the frequencies were used to estimate the parameters (gene frequencies and penetrances) of the best-fitting model for those regions where the data allowed it.

(3.1) *Data Analysis and Comparison*

The following quantities were specifically sought out:

Sex-specific frequencies: The frequency of MS in women and in men for various populations, or at least the ratio of these frequencies.

The risk to twins of MS patients.

The risks to other relatives of cases: Absolute or relative risks to siblings, parents, children, cousins, uncles, aunts.

The most difficult problem that must be faced in extracting frequency estimates from the data is that of adjusting for age, or allowing for mortality and residual risk.

The problem of variable diagnostic criteria (cf. Appendix 1) was also significant, but less critical, because it does not affect estimates for *ratios* of fre-

quencies, and these ratios are powerful tools for discriminating between hypotheses. However, the problem cannot be avoided when it is required to estimate penetrances.

The only familial study fully and scientifically age-adjusted was Sandovnick et al 1988. A useful result derivable from this study is a list of multipliers for residual risk to various categories of relatives of MS patients, The **residual risk multiplier** for a group G is the number m such that

lifetime risk to a random member of $G = m \times$ risk to date.

These multipliers (derived from Table I p. 536 and Table II p. 537) are as follows.

For a female patient:

mother: 1.01, father 1.04, parent: 1.01,
daughter: 3.84, son: n.a., child: 4.09,
sister: 1.37, brother: 1.39, sib: 1.38
aunt: 1.02, uncle: 1.03, aunt/uncle: 1.03,
niece: 4.7, nephew: 4.6, niece/nephew: 4.7,
maternal first cousin: 1.72,
paternal first cousin: 1.53
first cousin: 1.63.

For a male patient:

mother: 1.01, father: 1.01, parent: 1.01,
daughter: 5.7, son: n.a., child: 5.7,
sister: 1.3, brother: 1.35, sib: 1.34,
aunt: 1.01, uncle: 1.01, aunt/uncle: 1.0,
niece: 4.9, nephew: n.a., niece/nephew: 4.9,
maternal first cousin: 1.68,
paternal first cousin: 1.78,
first cousin: 1.74.

These multipliers are based upon the particular age-structure of the MS population and the general population in British Columbia, but since the course of MS is fairly uniform from country to country, it is to be expected that the multipliers may be used elsewhere, as long as the general standard of health care is compatible (poorer health care may lead to greater mortality from other causes, "saving" potential MS cases from developing the disease).

We note that the orders of magnitude of these multipliers make sense, if we consider the following facts:

The mean age of siblings (of cases who have siblings), and the mean age of first cousins, are both approximately the same as mean age of the cases. The mean age of parents, and the mean age of aunts or uncles are both about 25 years higher than the mean age of cases. The mean age of children is about 25 years lower than the mean age of cases.

The mean age of an MS case is in the middle forties.

Thus there is very little residual risk to parents, uncles or aunts, and multipliers of 1.30-1.38 for siblings and cousins and 4-5 for children look reasonable, since they agree with estimates of residual risk to persons in the middle forties and low twenties, respectively. The higher multipliers 1.53-1.78 obtained for cousins might indicate a slight bias in the data, involving an excess of cousins younger than the case. If anything, one would expect a slight excess of siblings younger than the case, since of two siblings the elder is more likely to have developed MS.

To understand the spread of sex-specific sibling multipliers, note first that the mean age of onset in men is about four years higher than in women. Thus the residual risk multiplier for men at a given age should be higher than that for women at the same age. The mean ages of male and of female sibs of cases should be the same, so this makes it reasonable that the multiplier for brothers of cases is higher than that for sisters of cases. A further factor is the fact that the average male case is older than the average female case, hence the average sib of a male case is older than the average sib of a female case, and hence has a lower residual risk multiplier.

These multipliers were used to make age-adjustments to all data, as required. A possible problem with this practice is that the above multipliers were derived from a study that confined attention to index cases with definite/probable MS, and was characterised by a relative strict application of diagnostic criteria (cf. Appendix 1). It is to be expected that definite/probable cases are, on average, *older* than possible cases (since the passage of time facilitates definite diagnosis), and hence the residual risk figures may be too low if applied to a study which included probable cases. It should be remarked that, in any case, the use of these crude methods of adjustment is only justified because the raw data are believed to suffer from comparatively large relative errors, of the order of 10-20% and upwards.

To convert prevalences to incidences, we used duration figures (cf. Appendix 2). In general, we take it that each such computation increases the relative s.d. by 10%.

To convert incidences to frequencies, we used the birthrate obtaining in the locality on the mean date of birth of cases. Again, we assume that each of these computations increases the relative s.d. by 10%.

Sex-specific frequencies

The ratio of the frequencies of MS in women and in men in a population will equal the ratio of the incidences in steady state, and as Matthews et al. observe, the estimates of that ratio in various populations range from 1.1 to 2.8.

It was important for us to attach realistic estimates of standard error to these estimates, and indeed to all frequency estimates.

For British Columbia, Sweeney et al. gave a prevalence of all types of MS of $130.5/10^5$ in people (sexes combined).[25] On the basis that the population is 2,780,0000, the s.d. was estimated at $2.2/10^5$. The mean age of cases was 44.7. From this the incidence/10^5 was estimated as

$$\frac{130.5}{2 \times (44.7 - 32)} = 5.14 \pm 0.6, \text{ per } 10^5$$

In using 32 as the notional age of diagnosis, we ignored the stated mean age of presentation (37), on the ground that this was probably biased upwards as one assumes that the neurologists files, on which the study was mainly based, are in many cases destroyed after some years. The s.d. was calculated on the assumption of a 10% error on the duration. Next, this incidence was divided between men and women, on the basis of the 1.95(±0.22) to 1 ratio obtained in the incidence study of Sadovnick et al.[23] In this, we were not ignoring the 2.18(±0.11) to 1 prevalence ratio obtained by Sweeney et al.[25] Our view is that these two ratios taken together pre-

dict a duration ratio of 1.12:1. In other words, in British columbia the duration in women is on average 1.12(±0.18) times as long as that in men. Finally, these incidences were converted to frequencies, assuming a birthrate of 2.43% (the birthrate in 1942, according to the Canada Yearbook 1985) with s.d. of 0.1%[32]. The result was a frequency of 0.00280±0.00098 in women and 0.00143±0.00050 in men. The ratio of the frequencies in women and men equals the ratio of the incidences, which is 1.95±0.22.

Here and elsewhere, in assigning s.d.'s to ratios, we used the principle based on the formula

$$\frac{a(1+\epsilon_1)}{b(1+\epsilon_2)} = \frac{a}{b}(1+\epsilon_1-\epsilon_2-\ldots)$$

that the absolute value of the relative error in a ratio is controlled by the sum of the absolute values of the relative errors in the numerator and denominator.

Ashitey and Millar 1970 give directly-measured mean frequencies of MS in people born in Northern Ireland in the period 1901-25 as 0.00121 in women and 0.00091 in men.[19] From Table I, p. 57, a relative s.d. of 15% was computed, and transferred to these data, giving 0.00121±0.00018 in women and 0.00091±0.00014 in men. The ratio of these frequencies is 1.33±0.43.

Schapira et al. 1963 quote a prevalence of $50/10^5$ in the population of Northumberland and Durham.[13] Assuming the population exceeds a million, we assigned a relative s.d. of 5%, making 50±2.5 per 10^5. From the quoted duration of 19.3 years, an incidence of 2.59±0.44 was computed. The propositi in the study were consecutive patients, so the sex-ratio in propositi is an estimate of the incidence ratio. The value is 1.27:1 or 0.56:0.44, and since there were only 700 cases, relative s.d.'s of about 20% are appropriate. Dividing the incidence according to this ratio, incidences of 2.90±1.17 in women and 2.28±0.92 in men were computed. Using the birthrate (1920-24) of 2.13%, quoted by Schapira et al. (p.322), frequencies of 0.00136±0.00055 in women and 0.00105±0.00042 were obtained.

Gudmundsson's incidence figures of 3.32 and 3.35 give an estimate of 0.99 for the sex-ratio in Iceland.[18] To compute s.d.'s, reference was made to his paper, Table 2, p.16, which lists the numbers incident in each of ten years. S.d.'s of 1.35 and 2.09 were calculated. The resulting s.d. on the sex-ratio was calculated as 1.27.

From birthrate tables, the rate for Iceland was 0.0285 in 1956, and this was the earliest value available to us. It was steady at about that until 1960, and than dropped. Assuming this rate, we calculated the frequency as

$$\frac{3.33 \times 10^{-5}}{2.85 \times 10^{-2}} = 1.168 \times 10^{-3}.$$

For the remaining studies, our immediate purpose did not require estimates of absolute frequencies, and only the sex ratios were estimated.

For Scotland (Shepherd and Downie),[12] the s.d. on the prevalence ratio of 1.49 was estimated at 5% relative, and using the British Columbia duration ratio, the resulting frequency ratio was 1.33±0.29. Using instead the duration ratio (1.38±0.14) calculated from the age-specific prevalence data and the notional diagnosis ages of 30 in women and 34 in men, the frequency ratio works out at 1.08±0.20, which is not significantly different.

On the basis of these examples, we infer that relative s.d.'s on sex-ratio estimates, however arrived at, will rarely be as low as 10%, and are typically around 20-30% in studies of regions in which the numbers of cases incident per annum is in two figures. Iceland, with a total population of only 160,000, and numbers incident per annum in low single figures, has a relative s.d. of 128%. Also, we observe that all these s.d.'s are large enough such that it does not make a significant difference whether sex ratio estimates are arrived at by one method rather than another.

Western Australia (McCall et al.):[4] The incidence is low, at $1.1/10^5$, and the number incident per annum was 6.6 on average. The s.d. on the sex ratio of 2.78 was estimated at a relative 80%, or 2.21.

In South Africa, the average number of white native cases incident per annum works out at $(0.9 \times 1389000 + 0.4 \times 1374000)/10^5 = 18$, and the estimated s.d. on the incidence ratio of 2.25 is 0.91.

The Faroes estimate of the sex-ratio need not be taken seriously at all, since the population is only in 5 figures.

An extraordinary prevalence ratio Kurtzke 1908 of 4:1 in the Sicilian city of Enna (15 cases, 12 female, 3 male) combined with a singularly high prevalence of $53/10^5$, should be mentioned.[14] We estimate the s.d. on this ratio as 1.6, so it does not differ significantly from 1, although it is certainly striking. The prevalence is significantly above that in other localities in Sicily.

We have noted that the gradient in the sex-ratio found by Kurtzke et al. 1979 is not statistically significant. The problem was the paucity of women in the U.S. armed forces.

The studies quoted by Gudmundsson show the following ratio of prevalences in women and men:[18] Sweden 1949: 1.49:1, Great Britain 1952: 1.79:1, England 1963: 1.43:1, Denmark 1956: 1.30:1, Scotland 1950: 1.35:1, Canada 1964: 2.31:1. We have no basis for attaching deviations to these figures.

In summary, the data on the relative risk to the sexes show the following:

1. The risk is in some places significantly greater in women, at the 99% level.

2. The risk is nowhere significantly greater in men, at the 95% level.

3. The female : male risk-ratio appears to vary from place to place. The hypothesis that the risk-ratio is constant at a value around 1.75 cannot be rejected at the 95% level on the basis of any single survey. The hypothesis that the risk-ratio is constant at any value can be rejected at the 95% level on the basis of the two best surveys, Sadovnick et al. and Schapira et al.[23,13] (χ^2=4.4 with 1 d.f.). However, it cannot be rejected on the basis of the set of seven surveys for which we have computed s.d.'s (χ^2=6.75 with 6 d.f.). The balance of the probabilities favours the variation, but constancy must be considered.

The risk to twins of MS patients:

To calculate risks from the data in Ebers et al. 1986 and Heltberg and Holm 1982,[2,31] we use the principle of multiple ascertainment (this was done in their analysis by the latter authors). Thus we use the fact that of 20+14=34 individu-

als with MS who had identical twins, 14 of those twins had MS, to estimate the MZ risk as 41.2%±8.4%. Similarly, of 44 individuals with MS who had DZ twins, 2 of those had MS, so the DZ risk was 4.45%±3.14%.

This does not differ significantly from the quoted risk of 1.9% to siblings in general. We made an age-adjustment, using a residual risk multiplier of 1.36±0.136, obtaining a risk of 56%±18% to MZ and 6.17%±5.31% to DZ.

In the Danish study quoted,[31] 4 of 19 MZ twins of cases were found to be concordant, and 1 of 28 DZ twins of cases. Thus the unadjusted risks are

8/23=34.7%±9.93 to MZ and 2/29=6.90±4.71 to DZ.

The age-adjusted risks are 47.2%±12% to MZ and 9.4%±8.0% to DZ.

The risks obtained in the two studies are not significantly different.

We ignored the other twin studies and the cases without clinical symptoms of MS.

Pooling the two studies, we obtained a risk of 51%±10% to MZ and 7.5%±4.4% to DZ.

Summarizing:

4. Pooling the studies, the risk to MZ twins of cases is significantly less than 80% and significantly greater than that to DZ twins. The risk to DZ is consistently found to be greater than to siblings in general, but the excess is not significant.

The risks to other relatives

First, we computed age-adjusted risks to parents and siblings, and the ratios of these risks, where these were not provided in the study.

Sadovnick et al. 1988 gave the adjusted risk to parents and sibs separately for patients of each sex.[23] The risks to parents and sibs of patients, without regard to sex, were calculated from the given risk and the female : male ratio of 1.95 in the index cases. The result was as follows: The risk to parents was 0.0267±0.0098, to sibs was 0.0391±0.0073, and the ratio sib risk/parent risk was 1.46±0.91.

Shapira et al. gave unadjusted risks.[13] Correcting these by using the residential risk multipliers derived above, we obtain the values in Table III. The adjusted sib risk was calculated as 0.01573±0.00614, the adjusted parent risk as 0.00586±0.00312 and the ratio 2.86±3.03.

TABLE III
North of England (age-adjusted)

frequency of MS in:	/10^5
women	136±55
men	105±42
sisters of female cases	2691±769
brothers of female cases	1175±524
sisters of male cases	1667±676
brothers of male cases	562±396
mothers of cases	837±373
fathers of cases	335±236

Applying the appropriate residual risk multipliers to the old data quoted in Spielman and Nathanson,[10] yielded the following recurrence risks and ratios:

Study	sib risk	parent risk	sib/parent
Germany			
1937	0.0012±0.0007	0.0±0.005	na.
England			
1951	0.0135±0.0050	0.0048±0.0017	2.8±2.3
Sweden			
1953	0.0106±0.0021	0.0033±0.0015	3.2±2.4

From Millar and Allison,[20] we observe that they found 14 parent-child pairs and 37 sibling pairs (counting three siblings in one family as three pairs) in Northern Ireland. The surveyed group consisted of 307 males with MS and 397 females, and the families of those patients (including other patients) included a total of 1500 male and 1405 female siblings, 2905 in all. Thus the unadjusted risk to sibs was 74/2905=0.0255±0.0029, which adjusts to 0.0349±0.0078. The unadjusted risk to parents was 0.0100±0.0027, and adjusts to 0.0101±0.0028. The ratio of sib risk to parent risk was 3.46±1.94.

In Orkney, Roberts et al. found concordance in 2 of 94 parents of cases,[7] and in 1 pair among 164 sibs of cases, from which we infer unadjusted risks of 0.0212±0.0150 to parents and 0.0122±0.0086 to sibs. The corresponding age-adjusted risk for parents is 0.0214±0.0151 and for sibs 0.0166±0.0134. The sib/parent risk ratio was 0.78±1.63.

Gudmundsson found 1 of 208 parents affected, and 3 sib pairs among 104 patients.[18] In the absence of family size data, we assumed a mean family size of 5, and obtained adjusted sib and parent risks of 2x3/416x1.36=0.0144±0.0142 and 1/208x1.01=0.0049±0.0049. The ratio is about 3±9.

In summary:

7. The average risks to sibs of cases is significantly less than 5%.

8. The risk to sibs is significantly larger than the risk to the general population.

9. The hypothesis that the risk to sibs is everywhere no greater than the risk to parents can be rejected. This is because all but one of many measurements of the risk show an excess risk to sibs.

Turning to sex-specific sibling and parent risks, these were given by Sadovnick et al. (Table II), and by Shapira et al. (Table I, adjusted in Table III). It was also possible to derive them from one further large-scale study, namely Millar and Allison 1954.[20] Using multiple ascertainment, the number of sisters of female cases found to be concordant in that study was 20, the number of brothers of female cases found concordant was 23, the number of sisters of male cases found concordant was 23, and the number of brothers of male cases found concordant was 8. Data were unavailable on the numbers of sibs of each sex of patients of each sex, and we assumed that the distribution of sibs by sex was independent of the sex of the patients. On that basis, the numbers of sibs were estimated as follows:

sisters of female cases:
1405×395/702=790.6,
brothers of female cases:
1500×395/702=844.0,
sisters of male cases:
1405×307/702=614.4,
brothers of male cases:
1500×307/702=656.0.

The unadjusted risks obtained were then:

to sisters of female cases:
20/790.6=0.02530±0.00558,
to brothers of female cases:
23/844.0=0.02725±0.00560,
to sisters of male cases:
23/614.4=0.03743±0.00766,
to brothers of male cases:
8/656.0=0.01220±0.00429.

Applying the risk multipliers of 1.37,

1.39, 1.3 and 1.35, respectively, these risks become:
to sisters of female cases:
0.03466±0.00764,
to brothers of female cases:
0.03788±0.00778,
to sisters of male cases:
0.04866±0.00996,
to brothers of male cases:
0.01647±0.00579

It is arguable that the s.d.'s here should be adjusted upwards by 10-20% to allow for error in the residual risk multipliers.

Seven fathers of cases had MS, and seven mothers (multiple ascertainment), so the estimated risks to fathers and mothers of cases are equal.. Age-adjusted, both risks work out at 0.0101±0.0040.

The facts about Northern Ireland are summarised in Table IV

TABLE IV
Northern Ireland (age-adjusted)

frequency of MS in:	/10^5
women	121±18
men	91±14
sisters of female cases	3466±764
brothers of female cases	3788±778
sisters of male cases	4866±996
brothers of male cases	1647±579
mothers of cases	1010±400
fathers of cases	1010±400

From data in Millar and Allison 1955 and Ashitey and Millar 1970.[20]

(3.2) Oligenic analysis

The following assumptions were made:

1. In order to develop MS, an individual must be of the susceptible genotype (have "genetic MS") and "be unlucky" in addition.

2. The relevant loci are di-allelic and pairwise unlinked. There is no selection and there is a steady-state, random-mating equilibrium.

Non-genetic factors were dealt with by assuming that the genetic factors act with reduced penetrance. The *penetrance*, π, for a given population was defined as the probability that an individual born into that population, and having "genetic MS", will prove to be an MS3 case, i.e., will be diagnosed in life as having clinically definite, probable, or possible MS.

It should be noted that this definition makes the penetrance depend upon standards of medical care in the population concerned. By defining the penetrance in this way, we relate it to a quantity which may be estimated statistically, and avoid such questions as how many 'symptomless' or 'sub-clinical' cases of MS may exist (see discussion above).

The hypothesis that the disease is fully non-genetic was also examined. This hypothesis is tantamount to the statement that genetic MS is a universal condition.We allowed for the possibility that the penetrance may be different in men and women for reasons unconnected with the loci under consideration (e.g. behavioural, hormonal, etc.).

We also allowed for the possibility that the environmental factor may be of such a kind that it is more likely to be shared by members of the population.

Hypotheses: The various hypotheses considered involved combinations of between zero to 4 of the following elements:

(A) an autosomsal locus,
(X) a locus on an X-chromosome,
(Y) a locus on the Y-chromosome,
(P) a reduced penetrance, π.

At the loci, the following required consideration:

(AD) autosonal dominant. We denote the frequency of the deviant dominant allele by *d*.

(AR) autosomal recessive. We denote the frequency of the deviant recessive allele by *r*.

(XD) dominant on X. We denote the frequency of the deviant dominant allele by *X*.

(XR) recessive on X. We denote the frequency of the deviant recessive allele by *x*.

(Y) As the disease occurs in women, we had only to consider a wild allele s^+ which protects against MS in men. We denote the frequency of the deviant allele 5 by y. (PS) unequal female and male penetrances, which we denote by π_f and π_m.

(F) a familial factor in the penetrance. If this hypothesis is active, we denote by Φ the probability that a member of the general population belongs to a vulnerable family, and by π the frequency of MS in genetically-susceptible members of vulnerable families. When we refer to F in the absence of P, it is understood thet π=1.

In order to save a great deal of space, we employed a shorthand for naming hypotheses. This is based on the labels AR, AD, XR, XD, Y, P, PS, and F introduced above. For instance, ARADPS (interpretred as AR+AD+P+S) denotes the following hypothesis:

(1) There are two loci involved in genetic MS, both autosomal, an AR locus (say m) and an AD locus (say s), and these loci are unlinked. To have genetic MS, an individual must have the MS type at both loci; at the AR locus, m, this means being homozygous m^-m^- for the MS allele m^-; at the AD locus, s, it means being homozygous s^-s^- or heterozygous s^+s^- for the MS allele s^-.

(2) To get MS, an individual must have genetic MS. But an individual may have genetic MS and not get MS. The probability , π_f that a woman with genetic MS develops MS may be different from the probability π_m that a man with genetic MS develops MS.

As a second example, ARPF denotes the following hypothesis:

(1) There is only one locus involved in genetic MS, an AR locus (say m). To have genetic MS, an individual must have the homozygous genotype m^-m^- at this locus, where m^- is the MS susceptibility allele.

(2) To get MS, an individual must have genetic MS, and must belong to a special class of susceptible families. Nothing is assumed known about the character of susceptible families, except the obvious corollary that all MS patients belong to susceptible families. The probability that an individual in the population belongs to a susceptible family is denoted ϕ. The probability that a member of a susceptible family who has genetic MS will develop MS is denoted by π. (It should be noted that under this hypothesis the usual concept of penetrance has to be modified; the probability that genetic MS will lead to overt disease in a member of the general population is the product $\phi \cdot \pi$, but the probability that genetic MS will produce MS in a person whose family has an MS case is simply π.)

We remark that in referring to MS susceptibility alleles as deviant, or de-

noting them m^- or s^-, etc., we do not intend to imply that these alleles are necessarily rare in the population. Multiple sclerosis is rare, but if it is the result of a combination of factors, then some of these factors may be common.

In testing hypotheses against the available data, we began with the simplest, and progressed to more complex ones, which involved a greater number of unknown variables, until a reasonable level of agreement was found. Our criterion of complexity in a hypothesis is the number of unknown variables, or adjustable parameters. The primary unknowns are the frequencies of the various deviant alleles in the gene pool, the penetrances, and the familial environment factor. For instance, ARADPS has four variables, namely the frequency r of the AR allele m^- among all m genes, the frequency d of the AD allele s^-, the female penetrance π_f, and the male penetrance π_m. Similarly, the hypothesis ARPF has three variables, the gene frequency r, the penetrance π, and the familial factor ϕ.

From each hypothesis were derived formulae to predict the probability that a brother, sister, sibling, parent, daughter, child, etc. of an MS patient will develop MS. In each case, the desired probability is the appropriate (unknown) penetrance times the probability that the relative in question has the MS genotype. The assumption that the loci are unlinked implies that the probability of having the genotype is the product of the individual probabilities that the realtive in question has the appropriate type at each locus, over the loci involved in the genotype. These individual probabilities are given by the standard formulae which apply to single locus inheritance (see below). Such formulae involve the (unknown) frequencies p_i of the troublesome allele in the population pool of genes at locus i.

For example, suppose we want to test the hypothesis ARAD, that MS is purely a genetic condition caused by two autosomal genes – one recessive, at locus m, and the other dominant, at a locus s unlinked to m. On this hypothesis, the predicted probability that a relative of a case will get MS is given by the probability that the relative has the genotype m^-m^- at locus m and carries the allele s^- at locus s. Since the loci are unlinked, this probability is the product of the individual probabilities that the relative has the appropriate type at each locus. To be specific, consider the risk to a sibling of a case. The case carries $m^-m^-s^-$, so the probability that the sib will carry m^-m^- is

$$^1/_4+^1/_2r+^1/_4r^2,$$

where r is the frequency of m^- in the pool of alleles at m, and the probability that the sib will carry s^- is

$$^1/_2+d-^1/_2d^2,$$

where d is the frequency of s^- in the pool at s. Thus the risk to the sib is the product

$$(^1/_4+^1/_2r+^1/_4r^2)\cdot(^1/_2+d-^1/_2d^2)$$

Similarly, the risk to parents of cases is the product

$$r\cdot(^1/_2+d-^1/_2d^2)$$

of the probability r of a parent having the AR condition and the probability $(^1/_2+d-^1/_2d^2)$ of a parent having the AD condition. The risk to members of the general population is the product $r^2\cdot(2d-d^2)$ of the risk r^2 of having the AR condition and the risk $2d-d^2$ of having the AD condition.

In some cases the formulas also involve the (known) sex-ratio for prevalence or incidence, as appropriate (–incidence ratios when the index cases are consecutive admissions, prevalence ratios when they are drawn from a national register on some day, etc.). For example, consider the hypothesis XRP, that MS is due to a recessive allele m^- at a locus m on X, acting with a reduced penetrance π that is independent of sex. Consider the predicted risk to the parent of an MS case, under this hypothesis. If the case is female, then the risk to the parents is $^1/_2(1+x)\pi$, where x denotes the frequency of the allele m^- in the pool, whereas if the case is male, then the risk to the parent is $^1/_2x\pi$. Letting α and β denote the fractions of the index cases that are, respectively, female and male, the risk to parents of index cases is thus the weighted average

$$(^1/_2(1+x)\propto+^1/_2x\beta)\pi.$$

Similarly, the risk to a sibling of a case is the weighted average

$$(^1/_4(1+x)\propto+^1/_4(1+x)^2\beta)\pi$$

of the risks to siblings of females and sibs of males.

The values of the fractions α and β for the three major familial studies we use are:

Sadovnick et al: α=0.661, β=0.339. Schapira et al: α=0.56. β=0.44. Millar and Allison: α=0.563, β=0.437.

Having obtained formulae for the predicted probabilities that various relatives are affected, we then compared these with the known data, and checked for concordance.

There are seven 1-variable hypotheses: AR, AD, XR, XD, Y, P, and F.

The 2-variable hypotheses include 15 two-locus hypotheses, such as ARAR(= two autosomal recessives), ARAD, ARXR, ARXD, ARY, ARP, ARF, ADAD, ADXR, etc., ten single-locus hypotheses, ARP, ARF, ADP, ADF, etc, and the non-genetic hypotheses PS and PF, a total of 27.

The 3-variable hypotheses include 35 three-locus hypotheses, ARARAR, ARARAD,..., XRXDY, 30 two-locus hypotheses, ARARP, ARADP, ..., XDYP, ARARF, ARADF, ..., XDYF, 10 single-locus hypotheses, ARPS, ..., YPS, ARPF, ..., YPF, and the nongenetic hypothesis PSF, for a total of 76.

The number of hypotheses grows very rapidly with the number of variables; their predictive powers decreases even more rapidly, since the data allow only limited discrimination between such hypotheses. From large list of four-variable hypotheses, we examined only some modifications of the more successful three-variable hypotheses.

The analysis was carried out in two stages. First, a preliminary, largely qualitative examination, based on the broad features of the data, ruled out many hypotheses. The more interesting hypotheses were then subjected to a quantitative analysis, comparing their predictions with the data from the three large familial studies.

4. Results

(4.1) Preliminary analysis:

(4.1.1) Analysis of one-variable hypotheses

It has long been known that no one-variable hypothesis will do. The purely genetic hypothesis – AD, AR, XD, XR, and Y – predict sibling frequencies of at least 25%, as opposed to measured values significantly under 5%, and they also predict 100% concordance in MZ

twins. The first two, the autosomal hypotheses, also fail to explain the sex-ratio in frequencies. The nongenetic hypothesis P predicts equal risks to siblings and the general public, and equal risks to men and women. The other nongenetic hypothesis, F, predicts equal risks to MZ and DZ twins, and equal risks to men and women. There are other objections to these hypotheses, but those given suffice.

(4.1.2) Analysis of two-variable hypotheses

In analysing a multi-variable hypothesis, the risks of concordance or recurrence on each factor must be multiplied together to obtain the risk of concordance/recurrence on the combination. Considering the factors separately, the recurrence risk to siblings of cases divided by the risk to parents equals:

(AD): $(^1/_2+d-^1/_2d^2)(^1/_2+d-^1/_2d^2)=1$

(AR): $(^1/_4+^1/_2r+^1/_4r^2)$ $r=(1+r)^2/4r$

(XD): $(^1/_2(1+^5/_4X-^1/_4X^2)\alpha+^1/_2(1+^3/_2X-^1/_2X^2)\beta)(^1/_2(1+^3/_2X-^1/_2X^2)\alpha+^1/_2(1+X)\beta)=(4+5X+X\alpha-X^2-X^2\beta)$ $(4+4X+2X\alpha-2X^2\alpha)$

(XR): $(^1/_4(1+x)\alpha+^1/_4(1+x)^2\beta)(^1/_2(1+x)\alpha+^1/_2x\beta)=^1/_2(1+x)(1+x\beta)/(\alpha+x)$

(Y): $(^1/_2(1+y)\alpha+\beta)(^1/_2(1+y)\alpha+\beta)=1$

(P): $\pi/\pi=1$

(PS): $(^1/_2\pi_m+^1/_2\pi_f)(^1/_2\pi_m+^1/_2\pi_f)=1$

(F): $\phi/\phi=1$.

Here a 1:1 sex-ratio for offspring is assumed, whereas α:β is the F:M sex-ratio of cases, with $\alpha+\beta=1$. Observe that α exceeds $^1/_2$, in practice, and from this it may be shown that the XD ratio is between 1 and 1.025, and the XR ratio is less than or equal to 1, no matter what value X has. Thus only AR loci have a noticeable effect on the sib/parent ratio (–we ignore the possibility that three or more XD loci are involved).

The relative female/male risks produced by the various factors are:

(AD): $2d-d^2/2d-d^2=1$

(AR): $r^2/r^2=1$

(XD): $2X-X^2/X=2-X$

(XR): $x^2/x=x$

(Y): $1/y$

(P): $\pi/\pi=1$

(PS): π_f/π_m

(F): $\phi/\phi=1$

To clarify the procedure, the sibling/parent risk ratio and the female/risk ratio for three sample hypotheses are:

(1) ARAD: sib/parent=$(1+r)^2/4d$, female/male=1.

(2) ADF: sib/parent = 1, female/male = 1.

(3) XDP: sib/parent = $(4 + 5X + X\alpha - X^2\beta)$ $(4+4X)+2X\alpha-2X2^2\alpha)$, female/male = $2-X$.

Proceeding to the analysis, the 15 two-locus hyopotheses have to be rejected, because they are purely genetic, hence 100% concordance in MZ twins. Of the 12 remaining, those five – ADF, ARF, XDF, XRF, and YF – which omit P (i.e. have $\pi=1$) fail because they predict that sibling risks are at least 25%. The nongenetic hypothesis PS fails because it has no familial factor, so it predicts that the risk to siblings equals the risk to the general poulation. The other nongenetic hypothesis PF predicts equal risks to the sexes. The hypotheses ADP and YP fall down on the relative sib/parent risk, because they predict it as 1. That leaves ARP, XDP, and XRP. Of these, only XDP predicts that MS is commoner in women. It predicts the F/M risk of MS as $2-X$. However, XDP predicts that the sib risk exceeds $^1/_2\pi$, half the penetrance, and the penetrance equals the risk to MZ twins. Since the data show that the risk to MZ twins is significantly greater than twice the risk to sibs, this hypothesis must also be rejected.

(4.1.3) Analysis of 3-variable hypotheses:

Of the 76 three-variable hypotheses, the 35 purely genetic ones may be rejected on the twin data. They all predict 100% concordance in MZ twins. Even allowing for deletion, it is inconceivable that it could operate frequently enough on one of three loci to reduce the concordance rate significantly. The 15 hypotheses involving two loci plus F also predict 100% concordance in MZ twins, and so are rejected. The nongenetic hypothesis PSF is also ruled out by the twin data, since it predicts equal risks to MZ twins and commonsex siblings. Of 125 hypotheses involving two loci plus P, only the nine containing an XD or a Y can explain the preponderance of female cases. Taking into account the sib/parent risk-ratio, and the MZ twin risk, the other locus would have to be an AR. This leaves 2 hypotheses in contention from this group: ARXDP, and ARYP.

Of 5 hypotheses involving one locus and PS, only ARPS and XRPS can explain the fact that the risk to Mz twins exceeds twice the risk to same-sex sibs, and only ARPS can explain the greater risk to siblings than to parents.

Of 5 hypotheses involving one locus and PF, only XDPF or YPF can explain the sex ratio, and these cannot explain the MZ twin/sib risk ratio or the sib/parent risk ratio.

This preliminary analysis of three-variable hypotheses leaves three hypotheses in contention: ARXDP, ARYP, and ARPS.

(4.1.4) Quantitative Analysis

We wrote computer programmes which selected, for each hypothesis input, the values of the parameters which minimise the sum (χ^2) of the squares of the normalised deviations in the predicted values of the eight data listed for Canada, England, and Northern Ireland in Tables II, III, and IV. This procedure is essentially equivalent to maximum likelihood estimation, simplified by ignoring covariances. Two types of algorithm were used: a systematic grid search, and a steepest-descent method (Appendix 3).

The results were as follows. In each case, we report the values of the adjustable parameter or parameters (variables) at the critical point and the value of the minimum χ^2. The appropriate number of degrees of freedom is 8 minus the number of variables.

Canadian data:

The best 1-variable hypothesis was P, giving $\pi=0.002$ and $\chi^2=65.6$. This was rejected (7d.f., $p<0.005$).

Among two-variable hypotheses, we computed:

PS: $\pi_f=0.0035$, $\pi_m=0.0015$, $\chi^2=62.16$

XRP: $x=0.034$, $\pi=0.047$, $\chi^2=41.1$

ARP: $r=0.129$, $\pi=0.111$, $\chi^2=12.997$

PF: $\pi=0.0305$, $\phi=0.056$, $\chi^2=12.927$

XDP: $X=0.02145$, $\pi=0.0667$, $\chi^2=2.401$

All but XDP were rejected (6 d.f., $P<0.05$).

We computed the 3-variable hypotheses ARPS, XRPS, XRXDP, ARYP,

XRYP, ARXDP, and ADPS. The most competitive were

ARXDP: r=0.305, X=0.0956, π=0.1608, χ^2=0.62026,

XRXDP: x=0.9055, X=0.023, π=0.0715, χ^2=2.298,

ARYP: r=0.148, y=0.48049, π=0.145, χ^2=4.407,

ADPS: d=0.0165, π_f=0.084, π_m=0.044, χ^2=5.1689,

ARPS: r=0.1395, π_f=0.1505, π_m=0.082, χ^2=6.27.

The remainder were rejected: XRYP has χ^2=16.753 and XRPS has χ^2=22.0587 (5 d.f., P<0.005).

We computed the following four-variable variations on these: ARXRPS, ADXRPS, ARYPF, ARXDPS, ARYPS, ARXDPF, ARXRPF and ADXRPF. The most competitive were:

ARXDPS: r=0.306, X=0.95, π_f=0.16, π_m=0.1615, χ^2=0.618,

ARXDPF: r=0.3055, X=0.0955, π=0.1605, ϕ=1.0, χ^2=0.62015,

ARYPF: r=0.2904, y=0.4785, π=0.1084, ϕ=0.316, χ^2=2.667,

ADXRPS: d=0.0245, x=0.6425, π_f=0.126, π_m=0.05, χ^2=3.636,

ARXRPS: r=0.171, x=0.6485, π_f=0.2125, π_m=0.089, χ^2=3.9195,

ARYPS: r=0.1465, y=0.5305, π_f=0.146, π_m=0.13, χ^2=4.2645.

The other two were rejected: ARXRPF had χ^2=11.044, and ADXRPF had χ^2=12.925 (4 d.f., P<0.05).

We note that the solutions for ARXDPS and ARXDPF coincide with that for ARXDP, in other words, no improvement in the fit of ARXDP is obtained by adding the assumption that penetrance varies with sex, or by adding the assumption that there is a familial environment factor. Also, ARYPS gives only slight improvement in fit over ARYP.

English data:

The computation of the same two-variable hypotheses gave:

ARP: r=0.1865, π=0.0315, χ^2=8.627,

XDP: X=0.0525, π=0.0155, χ^2=9.20754,

PF: π=0.007, ϕ=0.166, χ^2=12.4.

The hypotheses PS: χ^2=25.967, and XRP: χ^2=24.5 were rejected (6 d.f., P<0.005).

The computation of three-variable hypotheses gave

ARPS: r=0.1635, π_f=0.057, π_m=0.026, χ^2=3.1071,

XRPS: x=0.168, π_f=0.05, π_m=0.007, χ^2=5.716,

ARXDP: r=0.2285, X=0.392, π=0.042, χ^2=6.147,

ADPS: d=0.037, π_f=0.22, π_m=0.0105, χ^2=8.1788,

ARYP: r=0.199, y=0.7775, π=0.0325, χ^2=8.3194,

XRXDP: x=0.761, X=0.0675, π=0.019, χ^2=8.9583.

The hypothesis XRXY: χ^2=15.537, was rejected (5 d.f., P<0.01).

The four-variable hypotheses gave:

ARXRPS: r=0.2065, x=0.645, π_f=0.079, π_m=0.027, χ^2=2.3044,

ADXRPS: d=0.204, x=0.2525, π_f=0.0605, π_m=0.011, χ^2=4.9624.

The solution for ARXDPS has X=1, so reduces to ARPS, and similarly ARYPS reduces to ARPS, ARXDPF reduces to ARXDP, ARYPF reduces to ARYP, ARXRPF reduces to ARP, and ADXRPF has χ^2=12.4054, and was rejected (4 d.f., P<0.025).

Northern Irish data:

Of the two-variable hypotheses all but ARP were rejected (PF has χ^2=30.3, XDP has χ^2=40.089, PS has χ^2=83.77, and XRP has χ^2=89.366), and it has

ARP: r=0.1015, π=0.099, χ^2=12.1.

For three-variable hypotheses, we obtained

ARPS: r=0.101, π_f=0.1215, π_m=0.086, χ^2=7.0459,

ARXDP: r=0.108, X=0.719, π=0.11, χ^2=9.68,

and the others are rejected (ARYP: χ^2=11.965, ADPS: χ^2=28.295, XRXDP: χ^2=34.128, XRPS: χ^2=44.718, and XRYP: χ^2=44.756).

Of the four-variable hypotheses, none of ARXDPS, ARXRPS or ARYPS improve on ARPS, nor does ARXDPF improve on ARXDP, or ARYPF on ARYP. ARXRPF reduces to ARP (χ^2=12.099), and ADXRPF reduces to ADPF (χ^2=30.3), and both were rejected (4 d.f.). Finally, ADXRPS had χ^2=29.06 and was rejected.

As noted above, there is some evidence, based on pedigree analysis, for an MS susceptibility locus to the HLA system on chromosome 6, and for another close to the Ig system on chromosome 14.[10] The evidence marginally favours a dominant action for the MSS allele at the HLA-linked locus and a recessive action for the other MSS allele. In view of this, we felt it necessary to check the fit of hypotheses of the form ARAD + other factors. In view of the twin data, we must add P. At least one more variable is then added to explain the sex-variation. We examined ARADPS and ARADXDP. When ARADXDP was fitted to the familial data in Canada, England, and Northern Ireland, the results in all three cases was that the optimal value of d, the frequency of the AD allele, was 1; in other words, the addition of AD to ARXDP adds nothing to its explanatory power. In the case of ARADPS, much the same happened with the English and Northern Irish data, but with the Canadian data the optimal gene frequency involved a value of d less than 1. The solution was ARADPS: r=0.318, d=0.073, π_f=0.191, π_m=0.103, χ^2=3.85.

This compares with χ^2=6.27 for ARPS. In view of the loss of a degree of freedom, the reduction in χ^2 is not significant, but the hypothesis is not excluded.

5. Discussion

A total of hypotheses involving three or fewer variables, and a number of four-variable hypotheses were condsidered, to see how well they fit the data on MS.

The preliminary analysis led to the elimination of all one- and two-variable hypotheses, and the identification of ARXDP, ARYP and ARPS as the only 3-variable hypotheses capable of explaining the broad features of the data on the variation with sex, the occurrence in twins and the relative risk to parents and siblings.

The quantitative analysis led to the elimination, on the basis of one or another of the familial studies, of all two-variable hypotheses, although ARP performs better than any other. The analysis also led to the elimination of all three-variable hypotheses except ARPS and ARXDP. A number of four-variable refinements, ARXRPS, ARYPS, and ARXDPS also survived scrutiny, but the fit they give is not sufficiently

better than that of the simpler three-variable versions to warrant taking them seriously in the absence of data which might allow more refined discrimination.

The four-variable hypothesis AR-ADPS, for which segregation analysis involving genetic markers provides some evidence, was also not excluded, but predicts that the AD factor is universal in Britain and Northern Ireland and fairly rare in British Columbia. Given that most people in British Columbia are of British stock,[25] this seems improbable.

Thus ARPS and ARXDP were the leading contenders.

In attempting to decide which of these two hypotheses performs best, the following additional considerations were taken into account.

(1) On either hypothesis, the penetrance π or $^1/_2(\pi_f+\pi_m)$ is estimated by the MZ twin concordance rate, which is about 51%±10%. The sib risk is then predicted to be at least 6.4%±1.25% by ARXDP and 12%±2.5% by ARPS. The true value is estimated at 4% (from Sadovnick et al 1988). This tells against ARPS. However, the estimated penetrances from the quantitative analyses are all of the order of 0.05 to 0.20, so one could turn the argument around and argue that the data show that, on either hypothesis, MZ twins are about 3 to 10 times as likely to share the environmental factors as other pairs of people.

(2) ARXDP predicts that the sex-ratio of incidence between 1 and 2, depending on the frequency X of the deleterious XD allele. This prediction is consistent with the data. However, it implies that, other things (namely π and r) being equal, the frequency of MS will be higher in areas with lower sex-ratio. In view of the high measured sex-ratio and high MS frequency in British Columbia, and the low sex-ratio and high MS frequency in Britain and Northern Ireland, this implies that the penetrance is markedly greater in British Columbia, and/or the frequency of the AR allele is much greater. In fact, the quantitative studies say that π is much greater in British Columbia (16%) than in England (4%), and r is much greater in British Columbia (0.3) than in Northern Ireland (0.1), while π is not much greater in British Columbia than in Northern Ireland (0.11), and r is not much greater in British Columbia than in England (0.23). This says that genetically, British Columbia is more like North-east England than Northern Ireland, whereas environmentally British Columbia is more like Northern Ireland. The plausability of this is difficult to judge, since the nature of the environmental aspect is unknown. Sweeney et al remark that 59% of people in British Columbia are of British origin,[25] but there is a large component of British stock in Northern Ireland, too.

The other hypothesis ARPS implies that the penetrance sex-ratio π_f/π_m equals the incidence sex-ratio, hence varies from place to place. This would mean that the sex difference is not a function of physical female-male difference, but has something to do with a life-style sex difference which varies from place to place. This appears improbable. What do women do wrong in British Columbia, Western Australia or South Africa, as opposed to men in the same places or women in England?

(3) ARPS implies that the risk to siblings of cases should depend only on the sex of the sib, and not on the sex of the patient. Thus the risk to a sister of a female patient should equal the risk to a sister of a male patient, and the risk to a brother of a female patient should equal the risk to a brother of a male patient. This prediction, as far as brothers are concerned, is at variance with the Northern Ireland data in Table 4 (χ^2=4.98, 1d.f., P<0.05). In British Columbia and in England the difference in risk to brothers of male and female patients is not statistically significant.

ARXDP predicts that the ratio of the risk to brothers of female and male cases is

$$\frac{1+3X}{2+2X},$$

where X denotes the frequency of the XD allele. Given the estimated values of X in British Columbia (0.098), England (0.39) and Northern Ireland (0.69), this ratio should be 0.59 in British Columbia, 0.78 in England, and 0.91 in Northern Ireland. The measured values are 0.55 in British Columbia, 2.1 in England, and 2.2 in Northern Ireland. As with the other hypothesis, the discrepancy is significant in Northern Ireland.

(4) Unpublished data of Sadovnick (private communication) suggest that in multiplex pedigrees, apparent transmission of the disease is no less frequent through the unaffected father than the mother, i.e. paternal and maternal uncles have almost equal risks. For male and female index cases, the proportion of affected paternal and maternal uncles does not differ significantly.

This is consistent with ARPS, but ARXDP makes the following predictions: The ratio of risk to maternal and paternal uncles of female cases is

$$\frac{1+7X}{2+6X}.$$

For British Columbia, this gives a ratio of 0.65. The corresponding ratio for male cases is

$$\frac{1+3X}{4X},$$

giving a value of 3.3 in British Columbia. In the absence of the data, it is impossible to say whether or not these discrepancies are statistically significant.

From the pedigree data in Millar and Allison,[20] we extracted the fact that they found no paternal uncles or aunts, 2 maternal uncles of males, 2 maternal uncles of females, and 2 maternal aunts of females. The excess of maternal uncles of males over paternal uncles of males is predicted by ARXDP, and not by ARPS, but the excess of maternal uncles over maternal aunts and the excess of maternal uncles of females over paternal uncles of females is contrary to both hypotheses. The excesses are not statistically significant.

On balance, ARXDP was considered to be the best-fitting simple hypothesis. It explains the broad features of the population pattern of MS. It gives an outstanding fit to the Canadian familial data, which are the best data on MS in relatives. The only significant contrary evidence is the ratio of risk to brothers of female amd male index cases found in Northern Ireland by Millar and Allison 1954.

The hypothesis ARPS should also be taken seriously. There is more evidence contrary to it, but we do not regard it as conclusive, and more data would be

useful. For instance, it would be valuable to have sex-specific data on concordance in MZ twins.

Finally, we considered some alternative mechanisms lying outside the parameters of the foregoing study.

First, there is the possibility that 'genetic MS' is not a necessary precondition for MS, but increases susceptibility to the disease. Let α be the frequency of MS in the non-MS-genotype, and ß the frequency in the MS-genotype. Let r be the frequency of the MS-genotype. This hypothesis is only interesting if $\alpha(1-r)$ is roughly comparable to $ßr$, for otherwise essentially all or essentially none of the cases have the MS genotype, and we are back to cases considered above. But the fact that the relative risk to MZ and DZ twins of cases is 7:1 forces α to be an order of magnitude smaller than $ßr$, unless the environmental factor is at least 10 times more likely to be common to an MZ pair than to a DZ pair. Such a factor would almost have to be a mutagen, but typical mutagens would be thousands of times more likely to affect MZ. It should also be noted that migration data suggest that if one is prepared to ignore the twin data, then the incontrovertible data on increased risk in sibs tell us that α is at least 30 times smaller than ß, and the relative risks to parents and sibs tell us that we must still postulate an autosomal recessive. For sex variation this hypothesis predicts an interpolation between the predictions of the nongenetic hypothesis and some genetic hypothesis. Such interpolations cannot perform as well as the better extreme, which is the hypothesis isolated above.

Second, there is the possibility that the disease occurs principally in persons heterozygous at an autosomal locus, m, but in whom the wild allele m^+ has been deleted. Letting p denote the frequency of the allele m^-, and δ the frequency of the deletion of m^+, and treating the rest of the genotype as a contribution to the penetrance, π, we may assume that $p<0.1$, and we find that the frequency of the disease is roughly $2\delta p\pi$. There are two consequences. First, there must be selection to balance the deletion, for otherwise the homozygous double-deleted condition would have taken over the whole population. Opinions differ as to whether selection is significant in the case of MS, even when the condition is expressed. However, this is not a major point, because the selection needed to balance probable deletion rates would not be very substantial. The second consequence is that the disease will behave, at least as far as the contribution from locus m goes, essentially like an autosomal dominant condition acting with penetrance $\delta\pi$. But we have seen that there must then be *another* autosomal locus, at which the condition is recessive, and that the frequency of the condition at that other locus is commonly as low as about 1-2%. Thus $\pi<1/50$. In a country with an MS frequency of 1/1000, that would give $2\delta/50>1/1000$, so that $\delta>1/40$. This is an incredible deletion rate.

Third, there is the possibility of more than 2 alleles per locus. Such hypotheses introduce two extra variables for each extra allele, and rapidly lose all predictive power. However, one may rule out many of them by a slight elaboration of the methods used above. Alleles may be treated as identical unless they produce different penetrances. Working to the kind of accuracy appropriate to the data, there is not a great difference between a polyallelic locus m and a diallelic m' with a single deviant allele having penetrance equal to a weighted average of the non-zero penetrances at m.

Fourth, there is the polygenic threshold class of models.[7,21,10,16] The basic idea of these is that MS susceptibility is something like the length of a cow's tail, under the control of a great many genes, each contributing a certain weight, with susceptibility becoming significant at a certain critical loading. This hypothesis has a certain power. It is, however, more complicated than ARXDP, and it has the quantitative problem that consistent, and *consistently-different* values of heredita-bility have been derived from the data on siblings, on parents, and on cousins. From Roberts et al 1979, [7] we have this matched list of estimates (p.234, Table 2):

	Curtis '33	Pratt '51	Millar+Allison '54	Sutherland '56	Schapira '63
sibs	57±10	57±10	56±4	56±9	57±3
parents	44±18	40±11	48±6	46±15	43±8

This could be explained by supposing that sibs share an environmental factor to a greater extent than parents and others, but that assumption is not supported by the hereditability estimates in second and third degree relatives (cf. Roberts et al.).

6. Conclusions

It has long been known that no simple genetic mechanism can account for the population pattern of MS. In previous work on the problem, the combination of a single gene and reduced penetrance has been considered. Polygenetic models, involving many genes contributing linear effects and a threshold, have been considered, but simple oligogenic models have not been seriously studied in relation to the entire corpus of data. Efforts at genetic analysis have concentrated on pedigree studies. These are normally very successful, but the real power of pedigree studies is in searching for and identifying genetic marker loci. In a disease with low penetrance and of complex oligogenic type, studies of a few dozen cases are not very effective in determining the character of the inheritance.

We studied such models, using published population data, and we found as follows.

(1) No hypotheses having 2 or fewer variables provides a reasonable fit.

(2) Of 3-variable hypotheses, ARXDP is the best-fitting. This hypothesis states:

MS occurs only (or predominantly) in people who possess a genotype involving the combination of an autosomal recessive condition m^-m^- and a dominant s^- on X, and then only with a reduced penetrance.

It explains the broad features of the population pattern of MS, and gives an outstanding fit to the best data on MS in relatives.

Another three-variable hypothesis, ARPS, is less satisfactory, but could not be conclusively excluded. The hypothesis states:

MS occurs only (or predominantly) in people who are homozygous m^-m^- at an

autosomal locus m, and then only with reduced penetrance. The penetrance is greater in women.

(3) The hypothesis ARXDP yields the following conclusions:

1. The penetrance in Northern Ireland may be estimated at 11%, in North-east England at 4.2%, and in British Columbia at 16%. These compare with a value of 12% for Orkney obtained by Roberts et al. using segregation analysis.

2. The data do not support the inclusion of the hypothesis F that there is a strongly 'familial' environment factor, common to parents and siblings.

3. MZ twins do not share the environmental factor to a much greater extent than do DZ twins. Measured over predicted risks for MZ and for DZ twins of cases do not differ significantly.

4. The dangerous allele m^- at the autosomal locus m appears to be fairly uniformly distributed, occurring in for 10-30% of the gene pool.

5. There are large variatons in the frequency of the allele s^- at the X-locus s, ranging from negligible levels up to almost 100% of the pool.

6. The penetrance varies considerably with geographical location, but nowhere approaches 1, so that the environmental factor is of great importance.

(4) The hypothesis ARXDP makes possible the separate calculation of the environmental factor for each population, and may therefore help in attempts to identify that factor. It may also be used to compute risk tables to all categories of relatives in any population where sufficient data exist to permit the estimation of its three parameters.

Ebers (1983) called for a re-evaluation of the data on the geographic variation in MS incidence.[9] We believe that should now be done on the basis of ARXDP. The possible environmental factors should be correlated with the calculated penetrances, instead of the incidences or prevalences. It is also desirable pending the identification of the susceptibility loci and the development of diagnostic tests for the susceptibility alleles, to prepare software for calculating risks to relatives in each country on this theoretical basis. Such risk estimates would be more reliable than the empirical risks computated by Sadovnick et al.,[23] which are inevitably skewed by the genetic peculiarities of Western Canada, and cannot be used with confidence elsewhere, even if scaled in accordance with the local gross incidence.

The methods used here are capable of wide application, to other diseases and conditions, and other animals and plants. The full oligogenic calculus is a flexible tool which is capable, in theory, of discriminating among a wide range of genetic hypotheses, given sufficient population data.

APPENDIX 1:

The effect of variations in diagnostic criteria

From the studies cited, we assembled the following data on the division of cases between catgeories:

Millar and Allison 1954:[20] 476 probable, 145 possible, 79 early.

Schapira et al. 1963:[13] 24 reliable, 6 possible.

Dean 1967:[3] 158 probable, 21 possible.

Gudmundsson 1971:[18] 90 definite, 4 probable, 10 possible.

Brady et al 1979:[7] 1931 probable, 226 possible.

Roberts et al 1979:[7] 45 probable, 6 possible.

Shepherd and Downie 1980:[21] 324 probable, 193 early and latent probable.

Sweeney et al 1986:[25] 2596 definite/probable, 744 possible, 292 optic neuritis, 446 MS – a differential diagnosis, 26 diagnosis uncertain pending developments.

Elian and Dean 1987:[5] 20 probable, 2 possible.

To compare these, we grouped the categories definite and probable as class 1 and possible, early , optic neuritis, etc. as class 2. Expressing class 2 as a percentage of the total, we obtained (number of patients in brackets):

study	class 2 (%)	total number of patients
Millar	32.0%	700
Schapira	20.0%	30
Dean	11.7%	179
Gudmundsson	9.6%	104
Brady	10.5%	2157
Roberts	11.7%	51
Shepherd	37.3%	517
Sweeney	36.7%	4104
Elian	9.1%	22

We concluded that there is little consistency in the definite/probable versus possible classification in different studies.

From studies which report separately on the classification of men and women, we extracted the following:

	Sex	Class 1	Class 2	Total	% in Class 2
Dean	F	112	12	124	9.6
1967	M	46	9	55	16.4
Brady et al.	F	1113	129	1212	10.4
1977	M	818	97	915	10.6
Sweeney et al.	F	1769	729	2498	29.2
1988	M	827	307	1134	27.0

From this we note that the disparity is large only in the relatively small study, and is not statistically significant ($P>0.1$). A somewhat higher proportion of men might be expected to be classified as definite or probable, if were true (as has been suggested) that men are more reluctant to seek treatment – this might also partly account for the later reported onset in men.

We conclude that it is reasonable to assume that the choice and interpretation of diagnostic criteria in a study will not affect the *ratios* of prevalences, incidences, or frequencies, but will materially affect the absolute prevalences, incidences and frequencies themselves.

As far as absolute values are concerned, this aggravates the problem arising from the fact that the smaller the population studied, and the more persistently it is studied, the more MS is found.

APPENDIX 2:

The problem of durations

We are interested, because we wish to use it to convert prevalences to incidences. The quantity of interest is the duration of MS cases as diagnosed cases, i.e. the period from a diagnosis of possible/probable/definite MS until death. The first difficulty is that most studies report the duration from time of *onset*, i.e. *onset of symptoms*, whether or not the patient is examined at that time.

There is a pitfall here. Schapira et al. (1963) state that in a prevalence study the duration may be estimated as double the average time from onset to prevalence day, provided a steady state is

assumed.[13] This is correct only if one undersatands "onset" to mean "time of first diagnosis" (and one assumes that that is what Schapira et al. had in mind). However, if one doubles the average time from onset of symptoms to prevalence day, then one overestimates the mean duration of diagnosed MS by twice the average time from onset to diagnosis (and overestimates the mean time from onset to death by the average time from onset to diagnosis), and that can be a large error. In the worst case, in Gudmundsson's study,[18] the average time from onset to diagnosis was 12 years. Naturally, in the rule prevalence = incidence by duration, it is essential to understand "duration" to mean the duration of diagnosed MS, because prevalence and incidence are measures of the occurrence of diagnosed MS.

In order to be clear and concise, we will here use the word "term" to denote the period of time from onset to death, and reserve "duration" for the period from diagnosis to death.

Gudmundsson 1971 lists estimates of the average term from a number of previous studies.[18] These are as follows (country or city and date nominated, term in years): England (1917) 12, Germany (1933) 10, Denmark (1934) 10.3, England (1936) 12.5, Sweden (1942) 9.2, Sweden (1949) 13.0-34.0, USA (1950) 13.0, Boston (1950) Males 17.5 and Females 15.5, England and Wales (1950) 22.0, England (1955) 20.0, Denmark (1961) 30.0, Winnipeg (1964) 17.7, Israel (1964) 17.4, USA (Veterans) (1970) 35.

Gudmundsson 1971 provided the following data from his own study. The average term in 13 patients (9F, 4M) dead before prevalence day 1955 (and alive after 1946), was 16.7 in females and 23.2 in males (taking one male suicide out of account). The corresponding numbers for the 31 patients who had died by 1971 were (with s.d's attached) 26.4±2.9 in females, 18.7±4.1 in males, and 24.1±2.4 overall (Table 17, p.48). One has the extraordinary situation that 15 terms are below 19 years, 15 are above 28 years, and only one is in between. Also, the period of illness in those still living shows that these survival figures will need to be substantially revised (upward) in the end. The mean time from onset to prevalence day 1955 in the 91 (MS3) cases alive on that day was 13.9 years in females and 12.5 in males (p.41).

Schapira et al. 1963 quote durations of 19.5 in women and 19.1 in men.[13]

Kurtzke and Hyllested gave the average period from onset to prevalence day among Faroese patients alive on that day as 24 years in women and 25.7 in men.[26]

For English-speaking white South Africans, Dean gave a mean age of onset of 31, a mean time from onset to diagnosis of 8 years, and from diagnosis to prevalence day of 6 years.[3]

Shepherd and Downie gave the mean period from onset to prevalence day as 15.3 years.[21] They remarked that the rule of Schapira et al. gives a mean duration of 30.6 years.

Sweeney at al 1986 report that the mean period from diagnosis to death among those cases who died before prevalence day was 7.8 years.[25] The mean age of the cases alive on prevalence day was 44.7.

From these data we draw the following conclusions.

First, survival times have increased significantly over the course of the century.

Second, the terms and durations in those cases located but dead before prevalence day provide a very poor guide to terms and durations in general. In the nature of things, one would expect such a sample to provide an underestimate of mean duration, and in fact it may be very much below.

Third, the rule that duration equals twice the period from onset to prevalence day leads to absurd conclusions. In the extreme case of the Faroese data, it yields a mean duration of approximately 50 years. The more sensible rule that duration equals twice the period from diagnosis to prevalence day gives better results, but it has two problems. First, the estimates it produces show great variability, often attributable to relatively recent improvement in the level of diagnostic expertise in the locality studied. In the worst case, Gudmundsson's figures yield a mean duration of about 3 years when this method is used! It also suffers from the defect that, even if the period from onset to prevalence day is given, data on the period from onset to diagnosis are often lacking.

Arising from this, we propose the rule that mean duration is approximately twice the difference between the mean age of cases and the mean age of diagnosis. This rule is exact in a steady-state situation, and it has the advantage that the mean age of cases is usually reported, and the mean age of diagnosis may be fairly reliably estimated, if not given. In the absence of other information, or where specialist neurologist services are recent in the locality, one could use the rule that mean duration is approximately twice the excess between the mean age of cases and the notional mean diagnosis age of 32. If sex-specific figures are requried, one could use the notional mean ages of 32 for females and 34 for males. For instance, in the fairly common situation where an expert neurologist enters an area and carried out a comprehensive MS survey, the date of actual diagnosis will be close to prevalence day for many patients. To get a realistic duration estimate, one needs to estimate the age at which these cases *would* have been diagnosed by the expert, had he been attending to the area continuously, and from those we observe that mean figures of 30 in women and 34 in men are typical. A standard deviation of about 2 would seem to be appropriate on these figures, in view of the observed range of mean onset age in different populations (Matthews et al. Table 1.5, p.9).[1]

Fourth, where incidence and prevalence are both reported, the rule that prevalence = incidence by duration is reasonably exact, provided duration is estimated in the manner just specified. For Iceland, prevalence over incidence was 57.1/3.3=17.3. The mean age of patients (computed from Gudmundsson, Table 1, p.15) was 40, so twice 40-32 is 16. However, the rule can give bad results. For white South African natives prevalence over incidence was 9.6/0.9=10.6 in women and 4.1/0.4=10.25 in men, twice the mean time from diagnosis to prevalence day was 2x6=12 years, and twice mean age minus 32 was 2x(43–32)=22.

APPENDIX 3:
Computational methods and error analysis

The objective of the computations was to determine for each hypothesis the values of the adjustable parameters which gave the minimum value for χ^2, defined as the sum of 8 terms of the form

$$\left(\frac{\text{predicted value - actual value}}{\text{standard deviation}}\right)^2,$$

corresponding to the values of risk to men, women, mothers of cases, fathers of cases, sisters of female cases, brothers of female cases, sisters of male cases, and brothers of male cases. There were up to nine adjustable parameters, namely gene frequencies, penetrances, and the environmental factor. In our practice, no more than four parameters were floating in any one computation. All parameters ranged from 0 to 1.

Two kinds of programs were used, called "rough" and "fine". The programmes were written in the language MODULA-2, and were at first executed on a PC clone. Subsequently, they were transferred to a Micro VAX 3500.

Each program accepted as input the specification of a hypothesis and returned a set of values for the parameters and χ^2.

Rough programs input ranges and step-sizes for the adjustable parameters, and performed a straightforward grid search, computing the value of χ^2 at each grid point, and retaining the minimum found. They halted when the entire grid was searched.

Fine programs input starting values and step-sizes, and moved on the grid, adjusting only one parameter at each step, namely that parameter which produced the greatest decrease in χ^2. They halted when no significant reduction in χ^2 was obtainable by any such move.

Error analysis of the algorithm for computing χ^2 indicated that accumulated round-off in the computation was less than 50 times the single rounding error. Numerical experiments were consistent with this, indicating that, with 8-figure real arithmetic, accumulated round-off in the calculation of χ^2 was less than 10^{-5}, and movement of 2×10^{-5} in parameters produced variation in χ^2 that exceeded the round-off error, even near minima of χ^2. In running the fine programs, care was taken not to use step sizes so small that errors of computation might exceed the greatest change in χ^2.

Fine programs are capable of terminating in reasonable time with more accurate results than rough programs, but they will terminate at local minima. The procedure followed was to begin the investigation of each hypothesis with a rough program, using the smallest practical step-size, then input the results as the starting position of a fine program. The results were double-checked in a few cases by using a rough program with small step-sizes, at the limits of practical computation on the mVAX 3500 (100 million grid-points, about 3-4 hours c.p.u. time), and were found to be satisfactory. Incidentally, in no case was a second internal local minimum detected, but it did sometimes occur that there were several distinct minima, all but one on the boundary (i.e. some parameter equal to 0 or 1).

More sophisticated methods of searching for the minimum χ^2 were considered too vulnerable to error, in view of the well-known problems that arise in multivariate minimisation.

These results were presented in preliminary outline at the May 1988 meeting of the Irish Neurological Association.

References

1. Matthews, W. B., Acheson, E. D., Batchelor, J. R. and Weller, R. O. McAlpine's Multiple Sclerosis. Churchill Livingstone. Edinburgh, London, Melbourne, New York. 1985.
2. Ebers, G. C., Bulman, D. E., Sadovnick, A. D. et al. A population-based study of Multiple Sclerosis in twins. N. Engl. J. Med. 1986: 315, 1638-42.
3. Dean, G. Annual incidence, prevalence, and mortality of Multiple Sclerosis in white South-African-born and in white immigrants to South Africa. B.M.J. 1967: 2, 724-30.
4. McCall, M. G., Sutherland, J. M. and Acheson, E. D. The frequency of multiple sclerosis in Western Australia. Acta Neurol Scand. 1969: 45, 151-165.
5. Elian, M. and Dean, G. Multiple Sclerosis among the United Kingdom-born children of immigrants from West Indies. J. Neurology, Neurosurgery and Psychiatry 1987: 50, 327-32,
6. Makay, R. P. and Mirianthopoulos, N. C. Multiple Sclerosis in twins and their relatives. Arch. Neurol. 1966: 15, 449-462.
7. Roberts, D. F., Roberts, M. J. and Poskanzer, D. C. Genetic analysis of MS in Orkney. J. Epidemiol. Comm. Hlth. 1979: 33, 229-35.
8. Roberts, D. F. and Bates, D. The genetic contribution to Multiple Sclerosis. J. Neurolog. Sci. 1982: 54, 287-93.
9. Ebers, G. C. Genetic Factors in Multiple Sclerosis. Neurological Clinics 1983: 1, 645-654.
10. Spielman, R. S. and Nathanson, N. The genetics of susceptibility to Multiple Sclerosis. Epidemiological Reviews 1982: 4, 45-65.
11. Compston, A. Genetic factors in the aetiology of multiple sclerosis. In McDonald, W. I. and Silberberg, D. H. (eds): "Multiple Sclerosis". Toronto. Butterworths. 56-73.
12. Schumacher, G. A., Beebe, G., Kibler, R. F., Kurland, L. T., Kurtzke, J. F., McDowell, F, Nagler, B., Sibley, W. A., Tourtellotte, W. W. and Willman, T. L. Problems of the experimental trials of therapy in multiple sclerosis: Report by the panel on the evaluation of experimental trials of therapy in multiple sclerosis. Ann. N.Y. Acad, Sci. 1965: 122, 522-68.
13. Schapira, K., Poskanzer, D. C. and Miller, H. Familial and conjugal Multiple Sclerosis. Brain 1963: 86, 315-32.
14. Kurtzke, J. F. The geographic distribution of multiple sclerosis: an update with special reference to Europe and the Mediterranean region. Acta Neurolog. Scand. 1980: 62, 65-80.
15. Kurtzke, J. F. A reassessment of the distribution of multiple sclerosis. Part I. Part II. Acta Neurolog. Scand. 1975: 51, 110-36; 137-157.
16. Gonzalez-Scarano, F., Spielman, R. S and Nathanson, N. 19 Epidemiology. In: McDonald, W. I. and Silberberg, D. H. (eds): "Multiple Sclerosis". Butterworths. Toronto. 37-55.
17. Bauer, H, J. Multiple Sclerosis in Europe. J. Neurolog. 1987: 234, 195-206.
18. Gudmundsson, K. R. Clinical studies of MS in Iceland. A follow-up of previous survey and reappraisal. Acta Neurolog. Scand. 1971: Suppl. 48, vol. 47.
19. Ashitey, G. A. and Millar, J. H. D. Multiple Sclerosis in Northern Ireland. Ulster Med. J. 1970: 39, 55-63.
20. Millar, J. H. D. and Allison, R. S. Familial incidence of disseminated sclerosis in Northern Ireland. Ulster Med. J. 1954: 23 (Suppl. 2), 29-92.
21. Shepherd, D. I. and Downie, A. W. A further prevalence study of multiple sclerosis in north-east Scotland. J. Neurol. Neursurg. and Psychiatry 1980: 43, 310-5.
22. Roberts, D. F., Papiha S. S. and Poszkanzer, D. C. Polymorphisms and multiple sclerosis in Orkney. J. Epidemiol. Comm. Hlth. 1979: 33, 236-42.
23. Sadovnick, A. D., Baird, P. A. and Ward, R. H. Multiple Sclerosis: Updated risks for relatives. Amer. J. Med. Genetics 1988: 29, 533-41.

24. Sadovnick, A. D. and Baird, P. M. J. The familial nature of multiple sclerosis: Age-corrected empiric recurrence risks for children and siblings of patients. Neurology 1988: 38(6), 990-1.

25. Sweeney, V. P., Sadovnick, A. D. and Brandejs, V. Prevalence of Multiple Sclerosis in British Columbia. J. Canad. des Sci. Neurolog. 1986: 13, 47-51.

26. Kurtzke, J. F. and Hyllested, K. Multiple Sclerosis in the Faroe Islands: I. Clinical and epidemiological features. Ann. Neurol. 1979: 5, 6-21.

27. Kurtzke, J. F., Beebe, G. W. and Norman, J. E. Epidemiology of multiple sclerosis in US veterans: 1. Race, sex,and geographic distribution. Neurology 1979: 29, 1288-1235.

28. Brady, R., Dean, G., Secerbegovic, S., Secerbegovic, A-M. Multiple Sclerosis in the Republic of Ireland. J. Ir. Med. Assoc. 1977: 70, 500-506.

29. Allison, R. S. and Millar, J. H. D. Prevalence of disseminated sclerosis in Northern Ireland. Ulster Med. J. 1954: 23 (suppl. 2), 1-27.

30. Sadovnick, A. D. and MacLeod, P. J. M. The familial nature of multiple sclerosis: Empiric recurrence risks for first-, second-, and third-degree relatives. Neurosurgery 1981: 31, 1039-41.

31. Heltberg, A. and Holm, N. V. Concordance in twins and recurrence in siblings in multiple sclerosis. Letter. Lancet 1982: 1, pp. 1068.

32. Canada Year Book 1985. Minister of Supply and Services Canada 1985.