The Natural History of Multiple Sclerosis

B.G. Weinshenker and G.C. Ebers

ABSTRACT: Studies which have attempted to define the outcome of multiple sclerosis (MS) have methodologic difficulties arising from patient referral biases and the length of follow-up required, which make prospective studies of an inception cohort unrealistic. Means to improve the validity of retrospective natural history studies are suggested. Results of existing series are summarized and compared. Survival is only rarely shortened by MS, but disability to the point of requiring aids for ambulation occurs in 30-70% of patients by 15 years from onset of symptoms. Disagreement as to the percentage of patients who are ultimately bedridden by MS likely arises in large part due to differences in patient ascertainment and follow-up. The need to develop early clinical markers for the patient at high risk for rapid development of major disability is stressed.

METHODOLOGY

Studies of natural history should attempt to identify a cohort close to the onset of illness ("inception cohort") and provide longitudinal follow-up data over a duration appropriate to the expected time course over which the disease evolves. The requirements of such studies are discussed by Sackett. Several problems, some unique to MS, complicate study design: difficulties identifying patients from onset of MS; bias toward severe and unusual cases in tertiary centers that gather and report data on natural history; slow evolution of disease; loss of institutionalized patients to follow-up; uncertainty as to the optimum parameter to study so as to reflect outcome.

Sample Size & Ascertainment

A summary of the populations studied in existing series is presented in Table 1. Earlier studies were largely clinic based and included mixtures of inpatients and outpatients. S. Poser’s study clearly showed the expected bias toward rapid rate of progression of disability in a hospital based series compared to a community based series.

Prospective studies are unrealistic as patients often do not seek medical attention at the time of their first symptoms. It is almost inevitable that the most benign cases will be missed. Asymptomatic cases are found at autopsy and patients with a single attack are not reliably diagnosed. Patients are generally heterogenous with respect to duration of illness and disability when first evaluated at our clinic. Two strategies that minimize ascertainment bias are (1) separately evaluating patients seen...
## Table 1: Population Size & Patient Ascertainment

<table>
<thead>
<tr>
<th>Author</th>
<th>Location</th>
<th>Year</th>
<th>Population Size</th>
<th>Diagnostic Certainty</th>
<th>Ascertainment</th>
<th>Geographically Based</th>
<th>Mean Time From Onset to Ascertainment (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thompson</td>
<td>Dublin</td>
<td>1986</td>
<td>290</td>
<td>DP</td>
<td>H</td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Visscher</td>
<td>Los Angeles County</td>
<td>1984</td>
<td>941</td>
<td>DP</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detels</td>
<td>California</td>
<td>1982</td>
<td>886</td>
<td>DP</td>
<td>P</td>
<td></td>
<td>&lt;10</td>
</tr>
<tr>
<td>Clark</td>
<td>Pierce Counties</td>
<td>1982</td>
<td>834</td>
<td>DP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verjans</td>
<td>Belgium</td>
<td>1983</td>
<td>200</td>
<td>D</td>
<td></td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Patzold</td>
<td>Hanover, Germany</td>
<td>1982</td>
<td>102</td>
<td>NR</td>
<td></td>
<td></td>
<td>3**</td>
</tr>
<tr>
<td>Poser</td>
<td>Germany</td>
<td>1982</td>
<td>2058</td>
<td>DPP</td>
<td>+</td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Bromran</td>
<td>Gothenburg, Sweden</td>
<td>1981</td>
<td>312</td>
<td>DPP</td>
<td></td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Confavreux</td>
<td>Lyon, France</td>
<td>1980</td>
<td>349</td>
<td>DPP</td>
<td>+</td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Kurtzke</td>
<td>USA</td>
<td>1977</td>
<td>527</td>
<td>DP</td>
<td></td>
<td></td>
<td>3**</td>
</tr>
<tr>
<td>Percy</td>
<td>Rochester, Minnesota</td>
<td>1971</td>
<td>67</td>
<td>DP</td>
<td></td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Gudmundsson</td>
<td>Iceland</td>
<td>1971</td>
<td>104</td>
<td>DPP</td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>Fog</td>
<td>Copenhagen, Denmark</td>
<td>1970</td>
<td>73</td>
<td>D</td>
<td>+</td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Leibowitz</td>
<td>Israel</td>
<td>1970</td>
<td>266</td>
<td>DPP</td>
<td></td>
<td></td>
<td>11.5</td>
</tr>
<tr>
<td>Panelius</td>
<td>Turku District, Finland</td>
<td>1969</td>
<td>146</td>
<td>DP</td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1952</td>
<td>675</td>
<td>+</td>
<td></td>
<td></td>
<td>NR</td>
</tr>
</tbody>
</table>

* According to author, supported or unsupported, most cases in the area surveyed were identified

** Estimated from available data

H = Hospital; C = Clinic; P = Prevalence Survey; D = clinically definite; DP = clinically definite & probable; DPP = clinically definite, probable & possible; NR = not recorded

from onset (2) conducting geographically based studies, wherein all cases in a community are identified regardless of severity.

The first approach has been used by McAlpine,10 Kurtzke11 and by the UCLA group.12-14 The latter group restricted their study to patients with onset between 1960 and 1970, so as to avoid inclusion of those with long survival to the exclusion of others who had died before the prevalence date in 1970. Duration from onset to ascertainment in the series surveyed varies from 2-12 years. The longer the time the greater the potential bias as more severe or rare fatal cases could have been missed.

The second approach, ie. a complete community survey, has been used by several groups with varying degrees of verification. These investigators include the UCLA group for three U.S.A. west coast counties,12-14 Confavreux15 for Lyon, France, Leibowitz16-18 for Israel, and Gudmundsson19 for Iceland. Kurtzke's study of World War II veterans15 is unique in that the degree of ascertainment was high in the chosen sample, as compensation was provided for veterans with this diagnosis. The interval from onset to ascertainment was short. The study was limited, however, by the unavoidable restriction to otherwise healthy young males.

Several studies15,20 have included patients treated with immunosuppressive drugs which could have potentially altered outcome.

** Follow-up **

The key elements are as follows: 1) adequate duration 2) standardized recording of data 3) evaluation at a point of relative clinical stability rather than at the time of an exacerbation 4) adequate tracing of cases so that patients are not lost for reasons arising from the outcome of illness (death or institutionalization).

A prospective study of an inception cohort has not been done. In the few prospective studies12-14,20 patients are heterogeneous in duration of illness and disability at time of entry. In the remainder, follow-up can be classified as longitudinal15,19,21 or cross sectional.6,11,16,17,22,23 In the former, fluctuations unrelated to fixed disability generally do not have a major impact, while these can cause significant problems with interpretation in the latter, especially when data are extracted from traditional records. Kurtzke11 reported that of 34 patients with severe disability at diagnosis, 11.8% had only mild disability at 10 years and 32.4% had moderate disability. Such improvement
MRI may offer a sensitive and objective means of follow-up. Recent studies show a correlation between semi-quantitative assessment of lesion severity on MRI and the KDS. Others suggest that new lesions may develop without clinical correlates, challenging the significance of our clinical interpretations of disease stability. This is reconciled, in part, by the eloquence of the site involved by a plaque. The final concern of both patient and physician is undoubtedly disability, which is best quantitated at the bedside.

### Measures of Outcome

Measures of outcome must be sensitive yet clinically meaningful and must mirror fixed disability. As survival is only minimally altered by MS (see below), it is not useful as a primary measure of outcome. Similarly, employment status has a weak relationship with disability. Attack frequency is more dependent on such factors as age and duration than on outcome. Most investigators have not found a correlation between attack frequency and disability.

**Carefully recorded disability status at a point of relative clinical stability is the most desirable single measure. Several scales have been proposed, but the Kurtzke disability status (KDS) scale, which has recently been extended, is the most sensitive and useful. It is primarily a measure of mobility. It has often been divided by duration to generate a "progression index", thereby assuming linearity of the KDS. A frequency distribution of disability stratified according to duration is preferable given the ordinal nature of the scale.**
In Table 4 in terms of KDS equivalents. The disparity between the various series can, in part, be accounted for by differences in the scales used. At lower levels of disability (KDS 1-5), the Detels/Visscher, Hyllested and McAlpine scales are insensitive. Also, the KDS scale includes deaths due to MS (KDS 10) unlike the McAlpine and Hyllested criteria. This avoids the potential difficulty in series that show disability only in survivors.6,22 Paradoxically, the percentage of patients severely disabled declines with higher mortality.

Considerable disparity among the series is evident. McAlpine’s low estimate of the percentage of patients with disability equivalent to KDS 8 in his 1952 series5 likely reflects his clinic and hospital based ascertainment, which is prone to incomplete follow-up of patients with advanced disability. Kurtzke’s veteran series11 and Panelius’ population based series22 suggest that 29% and 14% of patients, respectively, followed 15 years from onset, are at or beyond the equivalent of KDS 8 (bedridden). With the exception of Panelius’ series, those cited suggest that 50-60% of MS patients 15 years from onset have not reached KDS 6 (aids required for ambulation).

With the degree of accuracy and precision that is necessary to plan efficient clinical trials, more precise data, expressed in terms of the expanded disability status scale of Kurtzke,26 are desirable.

### Attack or Relapse Rate
The relapse rate varies with age, being higher in the younger patient.20,24 It also depends on the duration of disease, decreasing with time from onset, and is dependent of the clinical course, whether progressive or stable.6,10,20,22,24 Calculation of relapse rate by averaging total number of relapses divided by patient years of a population at different stages of MS is, therefore, of limited value. Some accept nonspecific symptoms or brief fluctuations in neurologic dysfunction as attacks; others require that some degree of remission occur. Furthermore, retrospective data consistently show lower rates than prospective data.20,21 Comparisons among series are, therefore, difficult. The definition of an attack should adhere to established criteria.27,38 However, this definition is to some extent arbitrary. Broman24 distinguishes an episode of worsening disability, intermediate in rate of onset and duration between an attack and progression. He refers to such episodes, which can last months before stability is reached as “periods”. While this term may be meaningful, it is difficult to quantitate. Data on the attack rate from retrospective series, especially when gleaned from conventional medical charts, are of limited accuracy.

Representative published figures for attack rate are presented in Table 5. Variation of an entire order of magnitude from 0.14/year19 to 1.1/year20 exists. Differences are likely accounted for by differences in definitions, the prospective nature of the studies with the highest relapse rates and the frequency with which patients were assessed in the retrospective series. In Gudmundsson’s series,19 the low rate may well reflect the long duration from onset to ascertainment; in our experience, patients seen for the first time generally underestimate the frequency of their attacks in the early years of their disease.

### Parameters Predictive of Outcome

#### The Early Clinical Course
If the early clinical course of MS were predictive of later outcome, patients with an unfavourable prognosis could be better selected for study of therapeutic interventions which entail risk. The onset of progressively worsening disease is generally acknowledged to carry a poor prognosis;11,15,18,27,28 however, this point of general agreement is limited in practical application because transition from relapsing to progressive disease can only be determined after progression occurs. Kurtzke11 points out that “this is an observation after the fact and not a prognostic or predictive criterion”. It would be desirable to determine prognosis in the patient at a low level of disability. Two features that could be considered are the early development of disability and the early relapse rate.

Kurtzke11 claims that the disability status at 5 years from diagnosis is predictive of the disability status at 10 years and 15 years. Only 7.47% of those with mild disability at 5 years after onset were severely disabled (KDS 6-10) at 10 years after diagnosis and only 11.4% were severely disabled at 15 years. Given that 37% and 46% of all patients with duration 10 and 15 years respectively in Kurtzke’s series had progressed to KDS 6 or beyond (see Table 4), this provides a significant refinement in prognosis for the 20% who had only mild disability at 5 years. For the majority of patients with moderate disability (KDS 3-5) at 5 years, roughly 50% remained within that category at 15 years and 50% became worse with severe disability (KDS 6-10). Thus, the predictive value of moderate disability at 5 years which affected over 50% of cases was no better than inferences from the entire group.

![Table 4: Disability in MS as a Function of Duration](https://www.cambridge.org/core/core/terms)

<table>
<thead>
<tr>
<th>AUTHOR</th>
<th>K* P</th>
<th>M</th>
<th>C* K* P</th>
<th>M</th>
<th>C* K* P</th>
<th>M</th>
<th>C* K* P</th>
<th>M</th>
<th>C* K* P</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years From Onset</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>82</td>
<td>8</td>
<td>25</td>
<td>42</td>
<td>18</td>
<td>8</td>
<td>4</td>
<td>&lt;1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>77</td>
<td>25</td>
<td>37</td>
<td>68</td>
<td>32</td>
<td>17</td>
<td>9</td>
<td>&lt;1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>82</td>
<td>40</td>
<td>46</td>
<td>76</td>
<td>32</td>
<td>29</td>
<td>14</td>
<td>&lt;1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* In equivalents of the Kurtzke Disability Score (KDS) according to the conversion scheme of Detels.49
* Deaths due to MS (KDS 10) are included.

K = Confavreux;13 K = Kurtzke;11 P = Panelius;22 M = McAlpine

### Table 5: Attack Rate in MS

<table>
<thead>
<tr>
<th>Years From Onset</th>
<th>Attack Rate Per Person-Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Patzold20</td>
<td>1.8</td>
</tr>
<tr>
<td>Prospective</td>
<td>1.8</td>
</tr>
<tr>
<td>Gudmundsson19</td>
<td>0.35</td>
</tr>
<tr>
<td>Retrospective</td>
<td>0.39</td>
</tr>
<tr>
<td>Fog21</td>
<td>(first 5 years)</td>
</tr>
<tr>
<td>Prospective</td>
<td>0.39</td>
</tr>
<tr>
<td>Retrospective</td>
<td>0.39</td>
</tr>
<tr>
<td>Leibowitz26,17,18</td>
<td>0.39</td>
</tr>
<tr>
<td>Panelius22</td>
<td>0.26</td>
</tr>
<tr>
<td>McAlpine6</td>
<td>1.23</td>
</tr>
</tbody>
</table>
Detels\textsuperscript{14} concluded that the level of disability in a group of patients examined in 1972 predicted which patients would be worse when they were re-evaluated in 1979. However, the duration from onset of symptoms at the time of initial evaluation in 1972 varied by up to 10 years. Differences in the degree of progression were not stated, and the percentage of patients that worsened was no greater than twofold in the group with advanced disability compared to those with lesser degrees of disability. Once again, the percentage that worsened was greatest in patients with intermediate levels of disability (walking with aids), and of these, approximately 50\% were classified as being worse and 50\% were not.

The attack rate is, as previously noted, subject to variation in definition. While McAlpine\textsuperscript{10} suggests that a low relapse rate, particularly after the second year of disease, is associated with benign disease, this is not the experience of most investigators. Confavreux\textsuperscript{15} found that the mean duration between the first and second relapse is approximately 6 years in "benign" and "intermediate" cases (defined by the rate at which disability developed), compared to 0.9 years in "hyperacute" and 2.4 years in "acute" cases; however, the number of relapses was greater for benign than for more malignant MS. Thompson\textsuperscript{23} has also described a significant association between short first remission (less than 1 year) and increased risk of progressive disease. Both Patzold and Pocklington\textsuperscript{20} and Fog and Linnemann\textsuperscript{21} found no correlation between the rate of disease progression and relapse rate, although follow-up in the former was short, and in both studies progression was indexed only to the period of observation rather than to the entire course from onset.

Kurtzke\textsuperscript{11} found no association between attack frequency in the first five years and disability status at 10 or 15 years: however, the study was retrospective and the attack rate was predictably low.

In summary, it appears that minimal disability at 5 years from onset is clearly favourable, but intermediate disability at that time is of uncertain significance. No conclusion is possible as to the predictive value of the relapse rate until this parameter is rigorously defined and analyzed at a consistent point in a patient's course.

**Demographic and Clinical Features**

Several investigators have sought to determine if various clinical and demographic factors have prognostic significance (see Table 6). Chief among these have been the age of onset, sex, and initial symptoms. These parameters are not independent of one another; for example, the older patient often presents with pyramidal symptoms and pursues a more rapid course.\textsuperscript{39}

Many investigators fail to show a sex difference.\textsuperscript{12,15,23,27} Most find a worse prognosis in patients who are older at onset; however, the principal difference appears to be between patients less than or greater than 40 years at onset.\textsuperscript{12,18,28} Few show significant differences in outcome according to age at onset in patients younger than 40 years.\textsuperscript{12,18,28} with the exception of Thompson\textsuperscript{23} who recently reported a significant correlation between age and "benign" MS (KDS £3 at 10 years). There is poor agreement regarding the quantitative, if not qualitative prognostic value of initial clinical findings. Several authors find optic neuritis to be favou-

<table>
<thead>
<tr>
<th>Variable</th>
<th>Author</th>
<th>Relatively Favourable</th>
<th>No Influence</th>
<th>Relatively Unfavourable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Thompson\textsuperscript{23}</td>
<td>Female</td>
<td>+</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Visscher\textsuperscript{12}</td>
<td>Female</td>
<td>+</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Detels\textsuperscript{14}</td>
<td>Male</td>
<td>+</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Confavreux\textsuperscript{15}</td>
<td>Male</td>
<td>+</td>
<td>Male</td>
</tr>
<tr>
<td>Age At Onset (years)</td>
<td>Leibowitz\textsuperscript{18}</td>
<td>Female</td>
<td>&lt;40</td>
<td>&gt;40</td>
</tr>
<tr>
<td></td>
<td>Panelius\textsuperscript{22}</td>
<td>Female</td>
<td>&lt;31</td>
<td>&gt;31</td>
</tr>
<tr>
<td></td>
<td>Thompson\textsuperscript{22}</td>
<td>Young age</td>
<td>&lt;39</td>
<td>&gt;39</td>
</tr>
<tr>
<td></td>
<td>Visscher\textsuperscript{12}</td>
<td>&lt;20 vs 20-29</td>
<td>+</td>
<td>&gt;40</td>
</tr>
<tr>
<td></td>
<td>Verjans\textsuperscript{27}</td>
<td>29 vs 30-39</td>
<td>+</td>
<td>&gt;40</td>
</tr>
<tr>
<td></td>
<td>Posen\textsuperscript{28}</td>
<td>30-39</td>
<td>+</td>
<td>&gt;40</td>
</tr>
<tr>
<td>Initial Symptoms/Signs</td>
<td>Kurtzke\textsuperscript{11}</td>
<td>&lt;40</td>
<td>+</td>
<td>&gt;40</td>
</tr>
<tr>
<td></td>
<td>Leibowitz\textsuperscript{18}</td>
<td>Motor</td>
<td>Motor, Sensory</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Visscher\textsuperscript{12}</td>
<td>ON, Sensory</td>
<td>Motor, Sphincter</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gudmundsson\textsuperscript{19}</td>
<td>Motor, Sensory</td>
<td>Motor, Cerebellar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>McAlpine\textsuperscript{10}</td>
<td>Motor, Sensory, Coordination</td>
<td>Motor, Cerebellar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Posen\textsuperscript{28}</td>
<td>Motor, Sensory, Coordination</td>
<td>Motor, Cerebellar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Confavreux\textsuperscript{15}</td>
<td>Motor, Sensory, Coordination</td>
<td>Motor, Cerebellar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurtzke\textsuperscript{11}</td>
<td>Motor, Sensory, Coordination</td>
<td>Motor, Cerebellar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leibowitz\textsuperscript{18}</td>
<td>Motor, Sensory, Coordination</td>
<td>Motor, Cerebellar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fog\textsuperscript{21}</td>
<td>Motor, Sensory, Coordination</td>
<td>Motor, Cerebellar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gudmundsson\textsuperscript{19}</td>
<td>ON, Motor, Sensory, Coordination</td>
<td>Motor, Cerebellar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>McAlpine\textsuperscript{10}</td>
<td>ON, BS, Sensory</td>
<td>Motor, Motor + Sensory</td>
<td></td>
</tr>
</tbody>
</table>

\* According to Authors' Conclusions

ON = Optic Neuritis; BS = Brainstem

---

* According to Authors' Conclusions

ON = Optic Neuritis; BS = Brainstem
Pyramidal (motor) and cerebellar findings are generally acknowledged to be unfavourable. Some find that concomitant sensory symptoms have a favourable influence on prognosis compared to motor symptoms alone. Kurtzke found no difference in his veteran series depending on latitude or rural/urban residence at induction. Visscher, however, found a significantly worse outcome in patients, comparably ascertained, in Los Angeles County compared to those in King and Pierce counties in Washington state. This could be due to either environmental (ie. climate) or host (ie. genetic) differences.

The potential influence of HLA type on prognosis is addressed by several authors with discrepant conclusions.

**Conclusion**

Accepting inherent methodologic limitations, a consensus statement from available data follows. The outcome of MS is highly variable. The disease will most often pursue a remitting course initially, but progressive development of disability to moderate disability to severe disability over the next 10 years beyond the 50% risk noted above. Studies of the natural history of MS will continue to be confounded by difficulty identifying patients at onset and ensuring complete follow-up. Nonetheless, guidelines have been suggested against which methodology can be assessed to determine the reliability of data. Kurtzke’s veteran study provides important disability information on a uniquely ascertained group of patients seen over a long period of observation; however, the means of follow-up lack the accuracy of standardized longitudinal clinical information. The challenge for clinic based studies is, therefore, improvement in ascertainment and tracing of patients. With widespread use of experimental therapies, the natural history of MS may have to be determined from data collected to date.

**References**

10. McAlpine D. The benign form of multiple sclerosis. A study based on 241 cases seen within three years of onset and followed up until the tenth year or more of the disease. Brain 1961; 84: 186-203.