Clinical and Demographic Predictors of Long-term Disability in Patients With Relapsing-Remitting Multiple Sclerosis

A Systematic Review

Annette Langer-Gould, MS, MD; Rita A. Popat, PhD; Stella M. Huang, MS; Kristin Cobb, PhD; Paulo Fontoura, MD; Michael K. Gould, MS, MD; Lorene M. Nelson, PhD

Objective: To identify clinical and demographic factors associated with long-term disability in patients with relapsing-remitting multiple sclerosis.

Data Sources: We searched the MEDLINE (1966-May 2005), EMBASE, CINAHL, Cochrane, and PsycINFO computerized databases, and reviewed reference lists of retrieved articles.

Study Selection: We included studies that examined predictors of long-term disability in patients with relapsing-remitting multiple sclerosis. We excluded studies that did not distinguish relapsing-remitting multiple sclerosis from primary progressive multiple sclerosis, enrolled fewer than 40 subjects, observed subjects for less than 5 years, or collected follow-up information in less than 80% of the inception cohort.

Data Extraction: Two reviewers assessed study quality in 4 domains: cohort assembly, definitions and assessments of prognostic factors and outcomes, and statistical methods. One reviewer extracted data on the direction, magnitude, precision, and statistical significance of the effect of each predictor on prognosis.

Data Synthesis: Heterogeneity of study designs precluded us from pooling the results of 27 eligible studies. Study quality was limited by cross-sectional design, enrollment of prevalent cases from referral centers, and lack of multivariate adjustment. Sphincter symptoms at onset (hazard ratio, 1.1-3.1), incomplete recovery from the first attack (hazard ratio, 1.3-3.3), and a short interval between the first and second attack (hazard ratio, 1.6-1.9) were most strongly and consistently associated with poor prognosis. Other factors widely believed to be of prognostic importance, including sex and age at onset, demonstrated inconsistent or weak effects on prognosis.

Conclusions: The most robust predictors of long-term physical disability in relapsing-remitting multiple sclerosis are sphincter symptoms at onset and early disease course outcomes. These factors can be used to guide treatment decisions for drugs with significant toxicities.

Arch Neurol. 2006;63:1686-1691

MULTIPLE SCLEROSIS (MS) is a chronic inflammatory illness that usually begins in young adulthood. Two clinical subtypes of MS are distinguishable. Approximately 80% of patients have an initial disease course characterized by relapses and remissions. The remainder have primary progressive MS (PPMS) and experience progressive decline in neurological function from onset. To our knowledge, there are no effective therapies for PPMS, and almost all patients experience severe and early disability.

In patients with relapsing-remitting multiple sclerosis (RRMS), disability can result from 1 or more of the following: incomplete recoveries from relapses, development of secondary progressive MS (SPMS) (a delayed progressive course), or cognitive impairment. However, as many as 40% of patients with RRMS never develop a clinically important physical disability.1

Immunomodulatory therapies for RRMS can reduce the relapse rate and slow the progression of mild disability, but are also expensive, and some are potentially toxic.2,3 Furthermore, with the introduction of a highly sensitive diagnostic test—brain magnetic resonance imaging—the fraction of milder MS cases will likely increase.4 Clinicians need information about prognostic factors to help them distinguish between patients who are likely to develop disability—and, thus, should be treated—and those who are likely to have a more benign course and should not be exposed to potentially harmful therapies.
We systematically reviewed all English- and non–English-language studies of predictors of long-term disability in patients with RRMS.

**STUDY IDENTIFICATION**

We searched the MEDLINE, EMBASE, CINAHL, Cochrane, and PsycINFO databases in April 2004 and updated the MEDLINE search 1 year later. We also scanned reference lists of retrieved articles and relevant conference proceedings. Based on titles and abstracts, full reports were selected and evaluated for inclusion.

**STUDY SELECTION**

We included studies that (1) examined the effect of any clinical or demographic prognostic factor in patients with RRMS on physical or cognitive disability; (2) included at least 20 relapsing-onset participants in a cohort study or at least 20 relapsing-onset subjects with and 20 relapsing-onset subjects without significant disability in a cross-sectional or case-control study; and (3) for cohort studies, reported at least 5 years of longitudinal follow-up data for at least 80% of the inception cohort.

**STUDY QUALITY AND DATA ABSTRACTION AND SYNTHESIS**

We prospectively developed criteria to assess 4 different aspects of study quality (cohort assembly, measuring prognostic factors, measuring outcomes, and statistical methods) and applied them to included studies.

One investigator (A.L.-G.) abstracted relevant data from each study. Calculations to standardize data presentation were performed by 2 investigators (A.L.-G. and S.M.H.).

**RESULTS**

We identified 2812 potentially relevant titles. Eighty-nine reports met the inclusion criteria, including 21 unique studies5-32 of clinical and demographic characteristics. Sixteen reports5-15,17,18,20-25,32 examined multiple baseline characteristics and intermediate outcomes as potential predictors of long-term disability (Table). Eight studies10,24-30 focused on a single predictor (African American race,26,27 pregnancy,24,25 early age at onset, treatment with azathioprine,20 early disability,10 and later disease course disability18) and are not discussed further.

**STUDY CHARACTERISTICS**

In most studies, selection of subjects and potential prognostic factors seemed to have been driven by convenience rather than a priori hypotheses. Only 2 studies prospectively enrolled incident cases and were population based.5-7 Most studies were cross-sectional studies that enrolled prevalent cases from referral centers with varying years of disease duration and unknown periods of follow-up.

The characteristics of the cross-sectional studies8-17,18,21-23 that examined multiple predictors varied widely in sample size (range, 186-1844), but most had similar distributions of onset forms, proportions of patients with RRMS and SPMS, proportion of women, and mean age at onset.

**STUDY QUALITY**

As a whole, the 27 included published studies5-32 had acceptable methods for selecting participants, but methods were suboptimal for measurement of outcomes, measurement of prognostic factors, and analysis of data. Fifteen studies8 satisfied at least half of the criteria for cohort assembly, whereas only 9,6,7,13,15,20,22,31,7,5,7,13,15,20,27,30,32 and 5,10,12,20 studies satisfied at least half of the quality criteria for measurement of outcomes, measurement of prognostic factors, and statistical methods, respectively. Reporting of methods was insufficiently detailed in most studies. All but 2 studies relied on records obtained at routine medical care to assess prognostic variables and outcomes.

Eight studies6,8,9,12,27 reported the results of multivariate analyses predicting prognosis in patients with RRMS. These analyses were of varying quality. One study8 used a hypothesis-driven approach when choosing variables for the multivariate model, but 4 studies3,6,9,10 chose factors based solely on the P values from univariate analyses. Only 1 study11 systematically examined potential confounders and subgroup effects, such as relapse frequency and relapse symptoms.

**SPECIFIC PROGNOSTIC FACTORS**

The Figure summarizes the results by study. Heterogeneity in patient populations, definitions of prognostic factors, and definitions of disability precluded us from pooling results across studies. Differences in the direction or magnitude of effects between studies could not be consistently explained by cross-sectional vs cohort study design, definition of disability (Expanded Disability Status Scale score of 4 vs Expanded Disability Status Scale score of 6 vs SPMS), or effect estimate reported (hazard ratio vs odds ratio). However, changing the definition of disability altered the effect sizes derived from the same study populations for relapse frequency, age, multiple neuroanatomical regions involved, multiple neurological symptoms, and motor symptoms at onset. For age and optic neuritis at onset, varying the definition of the predictor significantly influenced effect sizes between studies.

**SEX**

Male sex is commonly believed to be a risk factor for poor prognosis in RRMS, but the evidence is mixed. Of the 10 studies that considered sex5,6,8-13,15,17 2 showed that men with RRMS were at significantly greater risk after adjusting for other factors5,11 and 3 showed a nonsignificant trend of increased risk among men.5,6,8 The remaining 5 studies5,12,13,15,17 found no effect for sex.

**AGE AT ONSET**

Older age at onset was associated with a worse prognosis in all but 1 study12 but the strength of the association varied depending on how older age at onset was defined and how disability was defined. When continuous measures of age at onset were used5,6,10,11,15 the

*References 5-15, 17, 18, 20-23, 25, 28, 30, 32.
risk of developing SPMS per decade ranged from 10% to 34%. When age was dichotomized, increased risk of disability for patients who were older at onset ranged from 9% to 192%. A single study reported a better prognosis in patients with an older age at onset when the cut point for older age at onset was set at 23 years.

Differences in the definition of disability also affected the results. Varying the definition of disability from development of SPMS to "severely impaired or lost walking" changed the odds ratio from 2.12 to 1.09 in the same population of patients. Thus, age at onset does not seem to be a robust predictor of disability.

### Table. Characteristics of Prognostic Studies Examining Multiple Predictors in Patients With RRMS

<table>
<thead>
<tr>
<th>Source</th>
<th>Country of Study Population</th>
<th>Enrollment Period</th>
<th>Population Based</th>
<th>Incident Cases</th>
<th>No. of Participants</th>
<th>Onset Forms, %</th>
<th>Disease Duration</th>
<th>Benign, RR, SP, EDSS Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runmarker and Andersen, 1993; and Eriksson et al, 2003</td>
<td>Sweden</td>
<td>1950-1964</td>
<td>Yes</td>
<td>Yes</td>
<td>255</td>
<td>79% 14%</td>
<td>&gt;25</td>
<td>NA 20 80 43</td>
</tr>
<tr>
<td>Amato et al, 1999; and Amato and Ponziani, 2000</td>
<td>Italy</td>
<td>1983-1990</td>
<td>No</td>
<td>Yes</td>
<td>224</td>
<td>85 15</td>
<td>9.8 ± 2.6</td>
<td>34 38 27 NA</td>
</tr>
<tr>
<td>Rudick et al, 2001</td>
<td>North America</td>
<td>1990-1993</td>
<td>No</td>
<td>No</td>
<td>160</td>
<td>100</td>
<td>0</td>
<td>14.3 ± 5.5</td>
</tr>
<tr>
<td>Confavreux et al, 2003</td>
<td>France</td>
<td>1957-1976</td>
<td>No</td>
<td>No</td>
<td>278</td>
<td>82 18</td>
<td>7.6 ± 7.0/13.3 ± 8.3</td>
<td>16 71 29 NA/NA</td>
</tr>
<tr>
<td>Bernardi et al, 1987</td>
<td>Italy</td>
<td>1978-1986</td>
<td>No</td>
<td>No</td>
<td>233</td>
<td>95</td>
<td>5</td>
<td>9.4 ± 7.1/9.9 ± 4.5</td>
</tr>
<tr>
<td>Thompson et al, 1986</td>
<td>Ireland</td>
<td>1980-1984</td>
<td>No</td>
<td>No</td>
<td>240</td>
<td>68</td>
<td>17</td>
<td>11.4</td>
</tr>
<tr>
<td>Rilse et al, 1992</td>
<td>Spain, Hungary, the Netherlands, Norway, and Germany</td>
<td>1986-1987</td>
<td>No</td>
<td>No</td>
<td>526</td>
<td>88</td>
<td>12</td>
<td>6.5</td>
</tr>
<tr>
<td>Trojano et al, 1995</td>
<td>Italy</td>
<td>1976-1991</td>
<td>No</td>
<td>No</td>
<td>309</td>
<td>81</td>
<td>19</td>
<td>7.3 ± 5.5/14.3 ± 6.6</td>
</tr>
<tr>
<td>Simone et al, 2002</td>
<td>Italy</td>
<td>1990-1991</td>
<td>No</td>
<td>No</td>
<td>793</td>
<td>86</td>
<td>14</td>
<td>9.6 (3-45)</td>
</tr>
<tr>
<td>Moreira et al, 2000</td>
<td>Brazil</td>
<td>1980-1997</td>
<td>No</td>
<td>No</td>
<td>261</td>
<td>86</td>
<td>14</td>
<td>&gt;8/8</td>
</tr>
<tr>
<td>Kantarci et al, 1998</td>
<td>Turkey</td>
<td>1994-1997</td>
<td>No</td>
<td>No</td>
<td>1259</td>
<td>88</td>
<td>12</td>
<td>7.3 ± 5.5/8.0 ± 6.0</td>
</tr>
<tr>
<td>Bergamaschi et al, 2001</td>
<td>Italy</td>
<td>NA-1997</td>
<td>No</td>
<td>Yes</td>
<td>186</td>
<td>100</td>
<td>0</td>
<td>7.3 ± 5.5/8.0 ± 6.0</td>
</tr>
</tbody>
</table>

Abbreviations: EDSS, Expanded Disability Status Scale; NA, data not available; RR, relapsing-remitting; RRMS, RR multiple sclerosis; SP, secondary progressive.

- Number of patients with clinically definite or probable MS, except for Myhr et al, who did not specify type of MS (definite, probable, or possible).
- Data are given as mean, mean ± SD, or mean (range).
- In cohort studies, these were outcomes among patients with RR onset; and in cross-sectional studies, these were subtypes of RR onset. In patients with relapsing-onset MS observed for equal lengths of time, development of an SP course or reaching an EDSS score of 6 (requiring unilateral assistance to ambulate) is a measure of a poor outcome. In cross-sectional studies, however, the distribution of these subtypes is influenced by referral center bias and unequal lengths of follow-up.
- Benign MS was defined as an EDSS score of less than 2.0 (abnormal neurological examination result, but normal ability to walk) after a 10- to 15-year disease duration.
- In cohort studies, data were given as percentage of patients with an EDSS score of 6 or more; and in cross-sectional studies, data were given as mean or mean ± SD EDSS score for patients with RR/SP MS.
- Of patients in Göteborg, Germany, 7% had a malignant-onset course; and in northeast Scotland, 2% of patients had an uncertain onset course.
- Data are given for patients with RR/SP MS.
- This study contained a population-based subgroup (n = 195 patients with clinically definite or probable MS), but the effect of prognostic factors in this subgroup was not reported separately.
- For the entire cohort (including those with primary progressive MS).
- Reported for all patients with relapsing-onset MS (did not distinguish between the RR and SP subtypes).
- Data are given as median (range) for patients with RR/SP MS.
Figure. Study-specific hazard ratios (HRs) and 95% confidence intervals (CIs) of long-term disability in patients with relapsing-onset multiple sclerosis for demographic and clinical predictors, including male sex (A), age at onset (B), motor symptoms (C), optic neuritis (D), sphincter symptoms (E), sensory symptoms (F), brainstem symptoms (G), cerebellar symptoms (H), incomplete recovery from the first attack (I), longer first to second attack interval (J), and early relapse frequency (K). Where CIs are missing, it is because studies did not report them. If a study reported only median time, an HR was calculated by dividing median times, but P values reported reflect the results of the study-reported log-rank test. Squares indicate unadjusted estimates and diamonds represent adjusted estimates. EDSS indicates Expanded Disability Status Scale; EDSS score of 4, moderate neurological disability but fully ambulatory without aid; EDSS score of 6, requiring unilateral assistance to ambulate (eg, a cane); EDSS score of 6 or more, cane or worse; EDSS score of 3, essentially restricted to a wheelchair; EDSS score of 4 or more, wheelchair or worse; NA, data not available; NS, not significant; SPMS, secondary progressive multiple sclerosis; asterisk, P ≥ 0.20; dagger, subgroup analysis of patients with more than 5 years’ disease duration; double dagger, subgroup analysis of patients with 15 or more years’ disease duration; section mark, odds ratio compared with patients with relapsing-remitting multiple sclerosis or patients with normal or mildly impaired walking (Posei’); parallel mark, P <.05; paragraph symbol, P <.001; and number sign, P <.01.
TYPE OF SYMPTOMS AT ONSET

Except for sphincter involvement, the studies showed mixed, weak, or no effect of different types of symptoms at onset on prognosis. Bowel and/or bladder involvement was associated with an unfavorable prognosis. In 4 studies, spinal cord symptoms at onset showed no effect on disability in 2 studies, increased risk of disability (HR, 1.84) in 1 study, and varied effect depending on the definition of disability in a fourth study. Motor/pyramidal symptoms increased the risk of disability in 2 studies, showed a trend toward an increased risk of SPMS, or showed no effect on disability. Sensory symptoms at onset had no beneficial or adverse effects on prognosis in 7 studies and marginally decreased the risk of developing SPMS (HR, 0.71; P=0.08) in 1 study. Brainstem symptoms at onset showed mixed results; 1 study reported decreased odds, and the 2 remaining studies reported no significant effect. Cerebellar symptoms at onset were associated with a 50% to 60% increased risk of developing SPMS or showed no effect on disability. Sensory symptoms at onset had no beneficial or adverse effects on prognosis in 7 studies and marginally decreased the risk of developing SPMS (HR, 0.71; P=0.08) in 1 study. Brainstem symptoms at onset showed mixed results; 1 study reported decreased odds, and the 2 remaining studies reported no significant effect. Cerebellar symptoms at onset were associated with a 50% to 60% increased risk of developing SPMS in 2 studies but had no effect in 3 other studies. Patients who had optic neuritis as a presenting symptom had a lower risk of developing SPMS in 2 studies (HR, 0.52; odds ratio, 0.47), while 6 studies found no effect. One study found a large increased risk of developing SPMS (HR, 2.56; 95% confidence interval, 1.52-4.32) in patients with optic neuritis at onset, but this study defined optic neuritis as severe vs mild or no visual impairment, whereas the other studies defined optic neuritis as either present or absent.

EARLY OUTCOMES AS PREDICTORS

Early disease outcomes were remarkably consistent predictors of future disability. Incomplete recovery from the first attack was a consistently strong predictor of a poor prognosis in 5 studies (HR, 1.3-3.3). A longer time to a second attack increased the risk of developing long-term disability despite variability in defining the attack interval and the definition of disability in 3 studies. Conversely stated, a shorter interval between the first and second attack confers a worse prognosis (HR, 1.6-1.9). In contrast, a higher early relapse frequency was not associated with poor prognosis and, within studies, the magnitude of effect was influenced by the definition of disability. Bergamaschi et al showed that patients with high rates of motor and sphincter relapses are at high risk of developing SPMS. Thus, while relapse frequency is modestly predictive, the type of relapse may be more important. One study found that incomplete remission after a relapse, cognitive symptoms, and more severe deficits 5 years after the onset of MS were predictive of long-term disability. Achirom et al reported that having an Expanded Disability Status Scale score of less than 2 after 1 year was predictive of having an Expanded Disability Status Scale score of less than 3 after 10 years.

This systematic review found that the strongest and most consistent predictors of long-term physical disability in patients with relapsing-onset MS are sphincter symptoms at onset and early disease course outcomes. Bladder or bowel symptoms at onset, incomplete recovery from the first attack, a short interval between the first and second attack, and early accumulation of disability should alert clinicians to a potentially worse disease course.

Many MS experts believe that female sex, younger age at onset, optic neuritis, and sensory symptoms at onset indicate a more favorable prognosis in patients with RRMS, whereas motor or cerebellar symptoms at onset predict a more severe course.

In this methodologically rigorous and systematic review, we show that many of these factors have no consistent influence (eg, optic neuritis), weak effects (eg, sex, age at onset, and cerebellar involvement), or no effect (eg, sensory symptoms) on prognosis. A critical review of the existing literature does not support using these factors to guide treatment decisions or predict prognosis for patients with RRMS.

Many clinical trials in RRMS routinely enroll patients with a normal neurological examination result, regardless of disease duration and the test drug safety profile. The enrollment criteria for these trials make no attempt to target patients at high risk of developing disability. While this may be acceptable for safe drugs, it seems unreasonable for drugs with serious or unknown toxicities. One woman who received natalizumab during one such recent trial died of progressive multifocal leukoencephalopathy. According to the findings of this systematic review, she had no significant risk factors for developing long-term disability, and would have been unlikely to benefit from therapy. Until other reliable indicators of prognosis are identified, we recommend that enrollment in clinical trials of drugs with unknown safety or efficacy profiles be restricted to patients with early accumulation of disability, incomplete recovery from a first attack, and/or bowel or bladder symptoms at onset.

Findings from this systematic review suggest that relapse frequency by itself is an inadequate predictor of long-term disability in RRMS. Because counting relapses does not distinguish between mild and severe relapses, it is not surprising that the severity and type of relapse may be more predictive than relapse frequency. Accordingly, clinical trials of MS therapies should use sustained disability as their primary outcome, rather than relapse frequency. We also suggest that clinical trialists report the extent and duration of disability at baseline to ensure that participants are balanced on these factors after randomization.

Some of the most commonly cited MS natural history studies did not meet our inclusion criteria because they did not distinguish between PPMS and RRMS. Prognosis...
differs considerably for patients with PPMS and patients with RRMS. Therefore, analyses that do not exclude patients with PPMS may obscure positive associations or identified spurious predictors of prognosis in patients with RRMS.

An important limitation of our review is that the quality of the primary studies was suboptimal. All of the studies used outdated diagnostic criteria, and none examined predictors of cognitive disability. There were few population-based studies, few cohort studies, and few studies that used multivariate models to estimate the independent effects of prognostic factors.

These limitations highlight the need for new prognostic studies that enroll incident cases diagnosed during an era in which brain magnetic resonance imaging is widely used, use a cohort design, and analyze data by using hypothesis-driven multivariate models.

Accepted for Publication: May 12, 2006.

Correspondence: Annette Langer-Gould, MS, MD, Stanford University School of Medicine, HRP Redwood Building, Room T202 MC 5405, Stanford, CA 94305 (annette1@stanford.edu).


Financial Disclosure: Dr Langer-Gould has received consulting fees from Genentech, Amgen, and Biogen Idec; and Dr Nelson has received consulting fees from Amgen. As of September 11, 2006, Dr Langer-Gould is a full-time employee of Genentech (Dr Gould).

Funding/Support: This study was supported by K23 grant N543207 from the National Institute of Neurological Disorders and Stroke, National Institutes of Health (Dr Langer-Gould); and Advanced Research Career Development Award RCD 99-023-2 from the VA Health Services Research and Development Service (Dr Gould).

Acknowledgment: We thank Chris Stave for his assistance with the literature searches; Jean Dominique Delacroix, PhD, Carolina Albers, PhD, Alexander Kleschevnikov, PhD, Damien Colas, PhD, Pavel Belichenko, PhD, and Chengbiao Wu, PhD, for reviewing the non-English-language articles; and Elizabeth Hoyte for her help in preparing the forest plots.

REFERENCES