White Matter Lesion Progression in LADIS
Frequency, Clinical Effects, and Sample Size Calculations

Reinhold Schmidt, MD; Andrea Berghold, PhD; Hanna Jokinen, PhD; Alida A. Gouw, MD, PhD; Wiesje M. van der Flier, PhD; Frederik Barkhof, MD, PhD; Philip Scheltens, MD; Katja Petrovic, Msc; Sofia Madureira, PsyD; Ana Verdelho, MD; Jose M. Ferro, MD, PhD; Gunhild Waldemar, MD; Anders Wallin, MD; Lars-Olof Wahlund, MD; Anna Poggesi, MD; Leonardo Pantoni, MD, PhD; Domenico Inzitari, MD; Franz Fazekas, MD; Timo Erkinjuntti, MD;
on behalf of the LADIS Study Group

Background and Purpose—White matter lesion (WML) progression has been advocated as a surrogate marker in intervention trials on cerebral small vessel disease. We assessed the rate of visually rated WML progression, studied correlations between lesion progression and cognition, and estimated sample sizes for clinical trials with pure WML progression vs combined WML progression-cognitive outcomes.

Methods—Those 394 participants of the Leukoaraiosis and Disability Study (LADIS) study with magnetic resonance imaging scanning at baseline and 3-year follow-up were analyzed. WML progression rating relied on the modified Rotterdam Progression Scale. The Vascular Dementia Assessment Scale global score and a composite score of specific executive function tests assessed longitudinal change in cognition. Sample size calculations were based on the assumption that treatment reduces WML progression by 1 grade on the Rotterdam Progression Scale.

Results—WML progression related to deterioration in cognitive functioning. This relationship was less pronounced in subjects with early confluent and confluent lesions. Consequently, studies in which the outcome is cognitive change resulting from treatment effects on lesion progression will need between 1809 subjects per treatment arm when using executive tests and up to 18 853 subjects when using the Vascular Dementia Assessment Scale score. Studies having WML progression as the sole outcome will need only 58 or 70 individuals per treatment arm.

Conclusions—WML progression is an interesting outcome for proof-of-concept studies in cerebral small vessel disease. If cognitive outcome measures are added to protocols, then sample size estimates increase substantially. Our data support the use of an executive test battery rather than the Vascular Dementia Assessment Scale as the primary cognitive outcome measure. (Stroke. 2012;43:2643-2647.)

Key Words: clinical trials ▪ leukoaraiosis ▪ magnetic resonance imaging ▪ vascular dementia ▪ white matter disease

There is convincing evidence for progression of age-related white matter lesions (WML) from numerous population-based, community-dwelling, and hospital-based studies.1–11 Lesion progression occurs particularly in subjects with confluent lesions at baseline and parallels cognitive decline.2,5,9,12,13 WML progression was considered as a potential surrogate marker in therapeutic trials of cerebral small vessel disease; sample size calculations for trials using change in WML volume as outcome measures are available.14 The use of magnetic resonance imaging (MRI) surrogate markers such as WML progression may be helpful in proof-of-concept studies, but it is also desirable to know if therapy-related slowing of WML progression translates into significant slowing of cognitive deterioration.

Volumetric measurements of WML change are still labor-intensive and time-consuming, and most still need user interference. Specifically designed visual rating scales facilitate assessment and their sensitivity and reliability have been determined by comparing visual rating results with volumetric data in identical subjects.1,15–17
The Leukoaraiosis and Disability Study (LADIS) used the visual Rotterdam Progression Scale (RPS) to determine the change of white matter abnormalities over the course of 3 years and found lesion progression of various degrees in 74% of study participants. We extend previous work by evaluating the association between different grades of WML progression on the RPS and on cognitive functioning and by providing sample size calculations for clinical studies applying visually rated WML progression as surrogate end points.

Subjects and Methods

Data were collected as part of the multicenter, multinational LADIS study. The rationale and design of LADIS have been described elsewhere. In short, 639 elderly subjects were enrolled in a hospital-based setting. Included subjects had to be between age 65 and 84 years, had to show WML on MRI of any degree, had no or only mild disability on the instrumental activities of daily living, cancer, or other relevant systemic diseases; severe unrelated neurological diseases; leukoencephalopathy of nonvascular origin (immunologic, demyelinating, metabolic, toxic, infectious); severe psychiatric disorders; inability to provide informed consent; and inability or refusal to undergo cerebral MRI. Subjects were selected and stratified for WML severity according to the categorization of the modified Fazekas scale into 3 severity classes. The cohort was followed-up for 3 years. Reasons for dropouts during the observational period were death for 43, withdrawal of consent for 73, and refusal to undergo follow-up examination or second MRI scan in 73 subjects. In 1 center (47 individuals), funding was lacking for rescanning. Moreover, scans from 7 participants were excluded because of absent fluid attenuation inversion recovery images or inappropriate slice positioning, and in 2 subjects WML progression could not be evaluated because of bilateral infarctions. The current study cohort thus consists of those 394 study participants who underwent a baseline and 3-year follow-up scan that could be evaluated for WML progression. Their mean age was 73.1 ± 5.0 years. There were 210 (53.3%) women and 184 (46.7%) men. As described previously, subjects who underwent a follow-up MRI were younger, had higher education, higher diastolic blood pressure, lower total cholesterol, higher low-density lipoprotein, triglyceride, and glucose levels than subjects who were not scanned at follow-up. We also reported elsewhere that baseline MRI characteristics were comparable between the groups with and without follow-up MRI.

MRI Scanning and WML Progression Rating

MRI scanning at baseline was performed on 1.5-T scanners in 10 centers and on a 0.5-T machine in 1 center. Repeat scanning after 3 years was uniformly performed at a field strength of 1.5 T with 3 new MRI scanners (all 1.5 T) used. Thus, there was a change in field strength of the scanner in only 1 center. The magnetic resonance protocol has been described previously. It was kept identical over the observational time in all centers. At follow-up, visual rating of WML progression was performed in a side-by-side fashion blinded to clinical data. WML progression was rated on fluid attenuation inversion recovery images according to the modified RPS, in which absence or presence of progression (0 or 1, respectively) was rated in 3 periventricular regions (frontal caps, occipital caps, bands), 4 subcortical white matter regions (frontal, parietal, occipital, temporal), basal ganglia, and infratentorial region. The total progression score ranges from 0 to 9. The intrarater reliability was determined on 20 randomly selected scans that were scored twice, and calculation of weighted Cohen kappa resulted in an excellent 0.81 agreement.

Neuropsychological Testing

The test battery applied to all LADIS participants has been described elsewhere. For the purpose of this study, we used the Vascular Dementia Assessment Scale (VADAS-cog) and a composite score of executive function testing. The VADAS-cog scale comprises the Alzheimer Disease Assessment Scale and 6 subtests covering the cognitive domains attention, working memory, executive functions, and mental speed, which are frequently involved in vascular cognitive impairment. The VADAS-cog was applied according to the original version except for using only 1 trial instead of the original 3 when assessing word recall and word recognition to avoid an overly long test administration. The scoring system was identical to that used in the original version. The 5 tests added to the Alzheimer Disease Assessment Scale are a delayed recall of a word list, digit cancellation, a maze task, the symbol-digit-modalities test, digit span backward test, and verbal fluency test (animal category). The scoring methods used for these tests have been previously described. The VADAS-cog total score ranges from 0 (no impairment) to 120 (maximal impairment). Four tests specifically measured executive functions. These included the Stroop color-word test, the trail-making test, the symbol-digit-modalities test, and a verbal fluency test. To calculate an executive function score, all test results were z-transformed. For Stroop color-word test and the trail-making test, we focused on working speed and thus calculated the differences between time color-word card minus time color card, and time trail-making test version B minus time trail-making test version A, respectively. These difference scores were inverted (higher score denotes better performance) and added to the symbol-digit-modalities test and verbal fluency test scores, and a mean z-score was then calculated.

Statistical Analysis

The RPS was evaluated for the whole sample and the 3 WML groups. Differences in progression between the 3 WML groups were analyzed by a $\chi^2$ test. The Jonckheere Terpstra test was used to investigate the association between the change on RPS and change on VADAS-cog and the executive function score (EFS) over the course of the 3-year follow-up period. Based on these results, we performed sample size calculations for various scenarios of future clinical trials. All sample size calculations assumed a power of 80% and a 2-sided significance level of 5%. Calculations were based on the assumption that a particular treatment could reduce WML progression on the RPS by 1 grade over the course of the 3-year follow-up period. The sample size needed to detect reduced WML progression as the sole outcome was calculated by the Wilcoxon test. To determine the effect of reduced WML progression by 1 RPS grade on VADAS-cog and on EFS, we calculated a weighted mean. The mean of these scores in each RPS grade was weighted with the percentage of cases that would be expected within each grade after treatment. The sample sizes for the differences of the changes on VADAS-cog and EFS were then determined using a $t$ test. All analyses were performed using PASW18 Statistics, and $P<0.05$ was considered significant. For sample size calculation, nQuery Advisor 7.0 was used.

Results

The frequency of WML progression on the RPS is shown in Table 1. Results are given for the total sample and subsets with different WML grades at baseline. Any progression was seen in 290 (73.6%) of the 394 study participants. One hundred two (57.3%) participants with punctate WML, 110 (89.4%) with early confluent WML, and 78 (83.9%) with confluent WML at baseline showed WML progression after 3 years ($P<0.0001$). Baseline WML grade influenced the magnitude of future WML progression. A change of ≥4 grades on the RPS was present in only 12 (6.7%) individuals with punctate WML, whereas this was seen in 32 (26.0%) subjects with early confluent and 29 (31.2%) with confluent abnormalities ($P<0.0001$).
functioning on VADAS-cog in the total cohort, whereas no association between WML progression rating and cognitive sample and in the subset of subjects with grade 2 and 3 WML EFS.

The magnitude of decline on VADAS/cog and H11002 VADAS-cog (Jonckheere Terpstra test; P=0.004). As can be seen from the Figure, the magnitude of decline on VADAS-cog increased in an almost linear fashion with increasing progression of WML. Similar associations existed for the most extreme data points not considering outliers. Increasing WML progression rating, the median of VADAS-cog change is 8.5 points (range, 23 to 39 points) on EFS resulting from the reduction of WML progression as the outcome measures. As can be seen from Table 3, low numbers of subjects are required per treatment arm in scenario 1 trials in which WML progression is the primary outcome measure. Such settings allow for a further lowered sample size when including individuals with grade 2 or 3 WML, only. For scenarios 2 and 3, in which the expected change in cognition resulting from treatment effects on WML progression is used as the outcome, 2599 or 18 853 subjects per treatment arm would be required for VADAS-cog and 1809 or 1988 for EFS. Sample size strongly depends on both baseline WML grade and applied cognitive tests. Trials that would include only subjects with grade 2 or 3 WML at baseline and that use the VADAS-cog as cognitive outcome would require a sample size of 18 53 subjects per treatment arm.

Discussion
We found that using WML progression as an outcome measure in intervention trials in cerebral small vessel disease would require considerably less subjects compared with trials including cognitive outcome measures. Our study used the RPS to rate WML change over a 3-year observational period. As compared with quantitative methods, this scale allows for visual, and thus less labor-intensive and time-consuming, assessment of WML progression with good correlation to volumetric measurements. Other than volumetric measurement of WML progression visual rating is less likely to be affected by technical upgrades of MRI machines and changes in scanning sequences over time. During the conduct of the current study, MRI scanners were renewed in 3 study centers, with 1 center switching to a magnet with higher field strength. Thus, the results of the present study cannot be extrapolated to more sophisticated techniques of measuring change in WML which, however, may suffer from other difficulties when applied in a multicenter setting.

In line with previous studies, we showed baseline WML grade to be the strongest predictor of future progression. Consequently, interventional studies aiming to slow WML progression might enrich their study populations for participants with early confluent and confluent abnormalities to reduce sample size. According to previous reports of the Austrian Stroke Prevention Study and current LADIS data, 3-year intervention trials applying modestly effective treatments in subjects with grade 2 or 3 WML at baseline will need relatively small samples ranging between 116 and 490 subjects to provide significant results.

Table 1. Extent of White Matter Lesion Progression at 3-Year Follow-Up on Visual Rating According to Baseline White Matter Lesion Grade

<table>
<thead>
<tr>
<th>Rotterdam Progression Scale Rating</th>
<th>Early Punctate (n=178)</th>
<th>Confluent (n=123)</th>
<th>Confluent (n=93)</th>
<th>Total* (n=394)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>76 (42.7%)</td>
<td>13 (10.6%)</td>
<td>15 (16.1%)</td>
<td>104 (26.4%)</td>
</tr>
<tr>
<td>1</td>
<td>48 (27.0%)</td>
<td>19 (15.4%)</td>
<td>17 (18.3%)</td>
<td>84 (21.3%)</td>
</tr>
<tr>
<td>2</td>
<td>26 (14.6%)</td>
<td>33 (26.8%)</td>
<td>18 (19.4%)</td>
<td>77 (19.5%)</td>
</tr>
<tr>
<td>3</td>
<td>16 (9.0%)</td>
<td>26 (21.1%)</td>
<td>14 (15.1%)</td>
<td>56 (14.2%)</td>
</tr>
<tr>
<td>4</td>
<td>6 (3.4%)</td>
<td>14 (11.4%)</td>
<td>11 (11.8%)</td>
<td>31 (7.9%)</td>
</tr>
<tr>
<td>5</td>
<td>4 (2.2%)</td>
<td>9 (7.3%)</td>
<td>9 (9.7%)</td>
<td>22 (5.6%)</td>
</tr>
<tr>
<td>6</td>
<td>2 (1.1%)</td>
<td>8 (6.5%)</td>
<td>7 (7.5%)</td>
<td>17 (4.3%)</td>
</tr>
<tr>
<td>7</td>
<td>1 (0.8%)</td>
<td>1 (1.1%)</td>
<td>2 (0.5%)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1 (1.1%)</td>
<td>1 (1.1%)</td>
<td>1 (0.3%)</td>
<td></td>
</tr>
</tbody>
</table>

RPS indicates Rotterdam Progression Scale.
*RPS rating for the total group was first reported in reference 18.

There were 346 (87.8%) individuals who underwent both baseline and 3-year testing on VADAS-cog. A total of 313 participants repeated executive function testing. Over the course of the observational period, study participants deteriorated by 2.6±8.5 points (range, −23 to 39 points) on VADAS-cog and −0.7±0.44 points on the EFS. As can be seen from the Figure, the magnitude of decline on VADAS-cog increased in an almost linear fashion with increasing progression of WML. Similar associations existed for the radar plot showing the 25th and 75th percentile values. Whiskers extend to the display by the central mark within the box. Box edges represent the most extreme data points not considering outliers. Increasing WML progression was significantly related to deterioration on the VADAS-cog (Jonckheere Terpstra test; P<0.004).
The situation changes dramatically if clinical effects related to WML progression are used as study outcome. As expected, there was a significant but only moderate relationship between lesion progression and decline in cognitive performance. Much to our surprise, this association became attenuated rather than enhanced when focusing on subjects with grade 2 or 3 lesions. This was caused by a pronounced cognitive deterioration in individuals with early confluent or confluent lesions at baseline, irrespective of change in lesion extent. Individuals with early confluent and confluent abnormalities at the initial examination but no lesion progression over the next 3 years experienced an even greater deterioration on VADAS-cog than those with the same lesion extent at baseline and mild to moderate WML progression during follow-up. Only those individuals with grade 2 or 3 lesions who later experienced ≥5 grades of progression had a cognitive decline beyond the value observed in subjects with no WML progression at all. Similar relationships also were observed for performance on executive function tests. These complex associations between longitudinal WML and cognitive change severely affect sample size estimates for trials using cognitive change resulting from treatment-related slowing of lesion progression as the primary outcome measure. Given that any treatment slows WML progression by 1 grade over a 3-year time period, such trials will need to be excessively larger than studies with WML change as the sole outcome. When including individuals with any WML grade, 2599 subjects would have to be recruited per treatment arm to detect a significant difference on VADAS-cog and 1809 subjects would have to be recruited for the composite EFS. For executive function, this estimate could be reduced if subjects with only grade 2 and 3 lesions were included, but with 18 853 participants per group the numbers would become extremely high for VADAS-cog. The LADIS sample likely reflects the patient population with WML encountered in everyday clinical practice, although it was stratified for lesion severity and did not include subjects with dementia. Thus, one cannot exclude that sample size calculations may not be fully applicable to trials focusing on dementia cases. Current sample size estimates are based on the simplifying assumption that a given treatment could decrease lesion progression by 1 grade consistently over all categories of WML progression. Such a scenario would be unusual, and thus, we can only provide approximate sample size estimates. Despite these uncertainties, such calculations are important because they show that the dimensions of clinical trials using WML progression ratings as surrogate outcomes need to change tremendously once investigators strive to demonstrate beneficial effects not only on WML progression but also on cognitive functioning.

### Sources of Funding
The study is supported by the European Union within the V European Framework Programme Quality of Life and Management of Living Resources (1998–2002), contract QLRT-2000-00446, as a concerted action.

### Disclosures
Dr Wallin is a member of the Speakers Bureau for Lundbeck and Pfizer. Dr Poggesi is a consultant and member of the Advisory Board for Stroke; Cerebrovascular Diseases, Int Journal Alzheimer Disease and Acta Neurologica Scandinavica. Dr Inzitari received a research grant from Bayer Italy for the investigation of brain white matter changes and honoraria for conferences from Bayer Italy and Boehringer Italy.

---

**Table 2. Association Between 3-Year White Matter Lesion Progression Rating and Subsequent Change in Performance on Vascular Dementia Assessment Scale or Executive Function Score**

<table>
<thead>
<tr>
<th>Rotterdam Progression Scale Rating</th>
<th>Change in VADAS-Cog</th>
<th>Change in Executive Function Score</th>
<th>Change in VADAS-Cog</th>
<th>Change in Executive Function Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean (SD)</td>
<td>P* Value</td>
<td>N</td>
</tr>
<tr>
<td>0</td>
<td>94</td>
<td>1.02 (7.35)</td>
<td>0.004</td>
<td>0.02 (0.34)</td>
</tr>
<tr>
<td>1</td>
<td>73</td>
<td>1.66 (7.35)</td>
<td>0.004</td>
<td>0.01 (0.34)</td>
</tr>
<tr>
<td>2</td>
<td>72</td>
<td>2.65 (10.10)</td>
<td>0.004</td>
<td>0.01 (0.34)</td>
</tr>
<tr>
<td>3</td>
<td>48</td>
<td>3.96 (7.68)</td>
<td>0.004</td>
<td>0.01 (0.34)</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>3.20 (10.31)</td>
<td>0.004</td>
<td>0.01 (0.34)</td>
</tr>
<tr>
<td>5</td>
<td>18</td>
<td>7.00 (8.45)</td>
<td>0.004</td>
<td>0.01 (0.34)</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>6.54 (10.10)</td>
<td>0.004</td>
<td>0.01 (0.34)</td>
</tr>
<tr>
<td>≥7</td>
<td>3</td>
<td>6.00 (4.36)</td>
<td>0.004</td>
<td>0.01 (0.34)</td>
</tr>
</tbody>
</table>

SD indicates standard deviation; VADAS-cog, Vascular Dementia Assessment Scale; WML, white matter lesion.

*Jonckheere Terpstra rest.

---

**Table 3. Sample Sizes per Treatment Arm According to Lesion Grade at Baseline in a 3-Year Trial in Which Treatment Reduces White Matter Lesion Progression by 1 grade**

<table>
<thead>
<tr>
<th>Trial Outcome Measures</th>
<th>All WML Grades</th>
<th>Early Confluent and Confluent WML</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduction of WML progress</td>
<td>70</td>
<td>58</td>
</tr>
<tr>
<td>VADAS-cog Change resulting from reduction of WML progression</td>
<td>2599</td>
<td>18 853</td>
</tr>
<tr>
<td>Executive function change resulting from reduction of WML progression</td>
<td>1809</td>
<td>1988</td>
</tr>
</tbody>
</table>

VADAS-cog indicates Vascular Dementia Assessment Scale; WML, white matter lesion.

*All calculations assumed a power of 0.80 and a two-sided significance level of 5%.
References

White Matter Lesion Progression in LADIS: Frequency, Clinical Effects, and Sample Size Calculations

on behalf of the LADIS Study Group

Stroke. 2012;43:2643-2647; originally published online August 9, 2012;
doi: 10.1161/STROKEAHA.112.662593

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2012 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://stroke.ahajournals.org/content/43/10/2643

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org/subscriptions/