Manuscript title: MD1003 (high dose Pharmaceutical grade Biotin) for the treatment of chronic visual loss related to optic neuritis in multiple sclerosis: A randomized, double-blind, placebo-controlled study

Journal: CNS Drugs

Authors: Ayman Tourbah, Olivier Gout, Alain Vighetto, Véronique Deburghgraeve, Jean Pelletier, Caroline Papeix, Christine Lebrun-Frenay, Pierre Labauge, David Brassat, Ahmed Toosy, David-Axel Laplaud, Olivier Outteryck, Thibault Moreau, Marc Debouverie, Pierre Clavelou, Olivier Heinzlef, Jérôme De Sèze, Gilles Defer, Frédéric Sedel, Carl Arndt

Corresponding author: Professor Ayman Tourbah, Department of Neurology and Faculty of Medicine, CHU de Reims, and UCRA, Reims, France, LPN EA 2027 Université Paris 8, Saint-Denis. Email: atourbah@chu-reims.fr

Online Resources 3. Efficacy assessments

Visual acuity (VA)

VA was measured using the Early Treatment Diabetic Retinopathy Study (ETDRS) LogMAR chart. The chart was presented with standard illumination. The letter stimuli were printed on a translucent panel and lit from behind. The patient read down the chart starting from the biggest letters until he or she reached a row where a minimum of three letters on a line could not be read. If a subject was reluctant to guess, he/she was encouraged to guess anyhow. Letters were not pointed at or presented in isolation, since the task of recognizing an isolated letter is different from the task of recognizing a letter in a letter chart format.

Scoring was done on a letter-by-letter basis: the total number of letters read correctly was counted. Each letter read correctly added one point to the score; each line added five points. Under this protocol three letters read on one line and two on the next line produced the same score as five on one line and none on the next.

The score corresponding to the maximum number of letters read was converted into logMAR units.

Visual field (VF)

VF analyses were performed using the standard automated perimetry method. The patient sat in front of a small concave dome in a small machine with a target in the center. The chin rested on the machine and the untested eye was covered. The patient was given a button to click on during the exam. The patient was set in front of the dome and asked to focus on the target at the center. A

computer then shone lights inside the dome and the patient clicked on the button whenever he saw a light, allowing the computer to automatically map and calculate the patient's VF.

Thresholds were reported in decibels (db), in a range of 0–50. Fifty db was the dimmest target the perimeter can project. On the other end of the scale, 0 db was the brightest illumination the perimeter could project. Automated perimetry gave normative data: values of sensitivity obtained in a given patient were compared to stored values that had been obtained from normal people in the same age group. The normal threshold (sensitivity) was defined as the mean threshold in normal people in a given age group at a given location in the VF.

The Total Deviation plot was a point-by-point difference of the patient's threshold from those expected in age corrected normal individuals. The Mean Deviation (MD) was derived from the total deviation plot. Like the total deviation plot, the mean deviation indicated any overall depression (or elevation) of the patient's hill of vision. A positive number indicated a better than normal field (elevation of the hill of vision). A negative number indicated a depression of hill of vision.

Pattern standard deviation (PSD) measured irregularity by summing the absolute value of the difference between the threshold value for each point and the average VF sensitivity at each point (equal to the normal value for each point + the MD). VFs with the age-normal sensitivity at each point had a PSD of 0, as did VFs in which each point was uniformly depressed from the age normal value. Thus, the largest PSD was registered for focal, deep VF defects.

Three different machines were used in the recruiting centers

- Humphrey Field Analyzer (HFA):
- Octopus 900
- Vision Monitor CV1 (Metrovision)

All of them were validated and able to measure the mean deviation VF defect which constituted the endpoint of the study.

In order to ensure the reproducibility that relied on a good cooperation of the patient, automated perimetry was performed once during the pre-inclusion visit and once at the inclusion visit.

Automated perimetry was performed at M0, M6 and M12. A quality control was performed by the core ophthalmic laboratory (Reims Hospital), taking into account the reproducibility of the examinations as

well as potential artifacts that could be detected form the initial recordings. These artifacts included (1) fixation losses of more than a certain percentage (the printout highlights this information) that would negate any comparison that the machine makes with stored normative data, (2) inappropriate fluctuations, (3) reproducibility of the optical correction used during the test. Fields with false-positive and false-negative errors may produce characteristic changes in the gray-scale print out with discrepancy between the gray-scale and calculated decibel values. The white scotomas in the grey scale print out might draw attention to high false-positive responses. In contrast, "clover leaf" pattern is characteristic of a fatigue field with high false negatives.

Visually evoked potentials (VEP)

To examine VEP pattern, the subjects were seated in a comfortable posture with their visual acuity corrected using trial lenses, and were instructed to keep staring at the center of the stimulus located at a 100 cm distance on a 20 × 30 cm black-and-white video display monitor.

The mean luminance of the checkerboard was 50 cd m^{-2} (40–60 cd m^{-2}) and contrast between black and white squares had to be high (equal to or greater than 80%).

The minimum analysis time (sweep duration) for all adult pattern-reversal VEPs was 250 ms post-stimulus.

To analyze both the pattern onset and offset responses elicited by onset/offset stimuli, the analysis time (sweep duration) had to be extended to 500 ms.

The monocular stimulation was standard. This required a light-tight opaque patch to be placed over the unstimulated eye. Care was taken to have the patient in a comfortable, well-supported position to minimize muscle and other artifacts.

A minimum of two recordings of each VEP condition was recommended to be acquired, measured, and displayed to confirm reproducibility of the data. VEP recording was repeated more times in cases, when the subject cooperation was poor.

The fixation stability of the eyes was monitored closely by an experienced electrophysiology technician.

The pattern-reversal VEP waveform consisted of N75, P100, and N135 peaks. These peaks were designated as negative and positive followed by the typical mean peak time. The amplitude of P100

was measured from the preceding N75 peak. The time from stimulus onset to the maximum positive excursion of the P100 was referred to as the P100 latency.

Two parameters were evaluated: (1) presence of a clear P100 wave, (2) P100 latency. These parameters were recorded in the electronic case report form (eCRF).

Then, selected values for latencies and amplitudes in the mean change analysis (among all values in the eCRF) followed the following rule:

- The selected latency had to be the one available on the lowest checkerboard
- If several values were entered (for instance for the two lobes), then the longest value at M0 had to be selected
- For the follow-up values (at M6 and M12), the values obtained with the same checkerboard and recorded at the same lobe had to be used
- Amplitude values had to correspond to the same eye/checkerboard/lobe/than the latencies
- If two latencies had the same value, then the lowest amplitude at M0 had to be selected.

To avoid any disparity of interpretation between different centers, a second reading was performed in a core ophthalmic laboratory (see quality control below). The core ophthalmic laboratory validated or not the value based on the quality control.

Improvement was defined as (1) the reappearance of a P100 wave not visible in a previous examination or (2) improvement of the P100 wave latency of at least 12 ms.

Optical coherence tomography (OCT)

Spectral Domain Optical Coherence Tomography (SDOCT) is a particular implementation of Fourier Domain OCT that collects all of the wavelengths of light at the same time using a specially designed spectrometer. This represents a technological improvement compared to Time Domain Optical Coherence Tomography (first, 2nd and 3rd generation OCT).

In the present protocol, only SDOCT was used (OCT 4th generation). Values of RNFL thickness and macular volume were collected at M0, M6 and M12. The quality of the print outs was controlled in the core ophthalmic laboratory and values entered in the eCRF were validated or not. A stabilization of RNFL and macular volume in patients treated by biotin compared to the placebo group was not

expected. Mean changes in RNFL thickness and macular volume in placebo and treated groups were compared at 6 months as well as during the extension phase.

Health outcome assessments

Clinical Global Impression Scale (CGI)

The Clinical Global Impression - Improvement (CGI-I) scale is a 7-point scale that requires the clinician to assess how much the patient's illness has improved or worsened relative to a baseline state at the beginning of the intervention and rate as: 1, very much improved; 2, much improved; 3, minimally improved; 4, no change; 5, minimally worse; 6, much worse; or 7, very much worse.

The CGI was assessed by the patient (subject global impression, SGI) and by the clinician (clinician global impression, CGI) at M6 and M12.

Multiple Sclerosis Quality of Life-54 (MSQOL-54)

The 14 sub-scores and 2 composite endpoints of the MS-QOL54 questionnaire were summarized by treatment group and visit between baseline and M6:

1. Change in health
2. Cognitive function
3. Emotional well-being
4. Energy
5. Health distress
6. Health perceptions
7. Overall quality of life
8. Pain
9. Physical health
10. Role limitations due to emotional problems
11. Role limitations due to physical problems
12. Satisfaction with sexual function

13. Sexual function

14. Social function

The National Eye Institute 25-Item Visual Function Questionnaire (NEIVFQ-25)

The NEIVFQ-25 or VFQ-25 consists of a base set of 25 vision targeted questions representing 11 vision related constructs, plus an additional single-item general health rating question (Mangione et al., 2001). It was validated in French. The VFQ-25 also includes an appendix of additional items that researchers can use to expand the scales up to 39 total items. These additional items have not been validated in French and were not used in the present study. The VFQ-25 takes approximately 10 minutes on average to administer in the interviewer format.

The VFQ-25 generates the following vision-targeted subscales: global vision rating (1), difficulty with near vision activities (3), difficulty with distance vision activities (3), limitations in social functioning due to vision (2), role limitations due to vision (2), dependency on others due to vision (3), mental health symptoms due to vision (4), driving difficulties (3), limitations with peripheral (1) and color vision (1), and ocular pain (2). Additionally, the VFQ-25 contains the single general health rating question which has been shown to be a robust predictor of future health and mortality in population-based studies.

References

Mangione CM, Lee PP, Gutierrez PR, et al. Development of the 25-item National Eye Institute Visual Function Questionnaire. Arch Ophthalmol. 2001;119:1050-8.