Conflicting Results on Ginkgo Research

P. Le Bars

Department of Psychiatry, New York University Medical Center

Although rare, publications of randomized clinical trials with negative findings could be of great importance to detect potential confounding factors that should be taken into consideration to estimate the broad efficacy of a treatment. This is especially the case when researchers face a ‘negative trial’ and have the opportunity to compare their methods and results with ‘positive’ studies previously published. However, such opportunities should be handled with the most objective (unbiased) and careful states of mind in order for the research to be as accurate as possible and able to reject or support previous hypotheses or possibilities. Those are the sine qua non conditions for successful explorations and useful conclusions, keeping in mind that, *stricto sensu*, trial outcomes should only be taken as guesses under a defined statistical risk and that comparisons between ‘positive and negative’ results should be cautious when the findings are drawn from studies with different settings and designs. Overall, failure to reproduce previous findings may not be as useful as an extensive report of adequate data that can support rational explanations of the trial mechanisms that may lead to one or the other outcomes.

To illustrate some of the bias in the interpretation of apparently negative results, the author will comment on the latest published work on ginkgo, i.e., Maastricht trial [1, 2]. In these two trial reports, one may be surprised by the number of imprecisions, inconsistencies, and even contradictions between ‘what was planned’, ‘what was done’, ‘what was reported’, and ‘what was discussed’. Considering the potential flaws of the study, the questionable comparisons made with previous ginkgo trials, and the untenable conclusions of these reports, to weigh the outcome results of this study, several points should be taken into account:

As a primary step, the characteristics of the studied groups should be considered. In this respect, the Maastricht trial population is unique for the Ginkgo research, as well as for the current research on cognitive disorders. As briefly mentioned in the paper discussion, the Maastricht researchers recognized that the inclusion criteria selected a special sample of a population living in a particular environment, i.e., assisted-living centers, and having very old age and high comorbidity. Al-

though not discussed, the differences with previously studied patient population [3, 4] go beyond these points when one scrutinizes the screening procedures applied to select the participants. Among others, the diagnosis was based on a historical case review, the differential diagnosis for dementia was not supported by full examinations (clinical and laboratory) at the time of the study entry, and the selection of participants showing so-called Age-Associated Memory Impairment (AAMI) was based on not-yet-tested criteria, e.g. Syndrome Kurz Test (SKT) cutoff points.

Since this trial aimed to study ginkgo in population with Alzheimer’s disease (AD) or vascular dementia (VaD), one should necessarily have included physicians to perform clinical screening examinations and differential diagnosis, further supported by concomitant laboratory and neuroimaging work up. Missing such important tools, one cannot be sure that the dementia sample included only AD or VaD. Furthermore, the trial concomitantly aimed to study the so-called AAMI, yet to be validated as a nosological entity. These objectives implicate the selection of a study population with a broad and heterogeneous spectrum of cognitive decline that could spread from borderline to severe and could be related to AD as well as any other medical conditions, yet unveiled at the study enrollment. Such composite population may show unpredictable response to study intervention leading to a dramatic increase of the study noise. Regardless of the validity of the AAMI concept, the Maastricht trial framework was so particular that one could question the applicability of its outcomes to any specific clinical entity, precluding secure comparisons with latest studies of ginkgo in AD and VaD.

Another aspect to be considered is the original and very complex design selected for the study, e.g., 12 stratification strata according to diagnosis, prognosis, and treatment, two successive double blind randomization phases, multivariate analysis exploring numerous potential ‘confounders’. At the term of such unconventional and ambitious endeavor, one may, for the least, regret that the relevance of the numerous exploratory factors tested was neither reported nor discussed! On the other hand, if ginkgo efficacy was the primary goal, it is strik-

ing that researchers targeted a 24-week endpoint, while they simultaneously included a double blind withdrawal phase that was expected, by patients and raters, to occur anytime after 12 weeks of treatment! Such procedure has never been tested before (withdrawal phases usually occur after the efficacy endpoint) and could induce uncontrollable flaws with a high potential to obscure a treatment effect, particularly during the withdrawal phase of the second portion of the trial. Therefore, if the ginkgo effect, per se, is still to be assessed within such a study design, efficacy should first be considered at a 12-week endpoint (at the end of the conventional first double blind phase), leaving the changes occurring after the middle of the trial to exploratory research. Unfortunately the sample size was not computed for allowing the detection of a potential effect after a so short treatment period, i.e., 12 weeks, thus dramatically weakening power and statistical significance at this early endpoint.

Regardless of the endpoints selected, further concerns on study power arise when one notices that the theoretical cutoff points, set on the primary outcome measures, were based on clinically relevant milestones in relation with the average deterioration expected during a half-year course of AD, e.g. 3 SKT, 4 NAA points. Such landmarks may be valid to study a responder rate after a 6-month treatment period in a pure dementia population, as was done in the Kanowski et al. trial [3], but are probably inadequate for a composite population with unknown evolution such as the sample included in the Maastricht trial. Focusing only on the dementia stratum would not improve the situation, since power was conditioned by enrolling 100 patients receiving ginkgo and 50 receiving placebo, numbers never reached at any endpoint in the dementia group (i.e., unfortunately less than 70 dementia patients were included in the whole study).

Turning to the largest group, which included the so-called AAMI participants, one faces another challenge regarding the selection of the primary outcome measurements. On the one hand, in view of the AAMI study definition (i.e., cognitive impairment not associated with dementia), one may assume that this heterogeneous group showed only minor social dysfunction, if any. Thus, dealing with such a preserved function, further in a non-challenging environment (i.e., assisted-living center), the assessment of daily living (i.e. NAA) may not have been the most sensitive tool to detect a treatment effect, particularly after 12 weeks of observation. On the other hand, since AAMI is not a clinical entity, the rating of the Clinical Global Impression (CGI) of this group could not be based on well-defined clinical probes. In this condition, one may assume that a noticeable portion of the clinical impression was drawn from subjective changes reported by the AAMI participants, reducing the relevance of the CGI.

Regarding memory impairment, none of the assessments reported in the publication purely measured memory, e.g. the NAA word list included only one word presentation, precluding any learning and relevant storage. While the SKT included

a delayed recall of memorized pictures and could be considered the primary outcome measurement for this field, the results were strikingly not presented in the publication and its memory factor most likely not analyzed separately. In such ambiguous situation, no solid conclusion can be drawn on negative or positive effect of ginkgo on the memory function of the AAMI group.

Taking into account all aforementioned limitations and numerous others that cannot be delineated in such a short space, one may not be able to gather any valid idea on ginkgo effects at all in either studied group. The controversial design of the Maastricht trial reflects the pilot character of this study. It may have been appropriate for epidemiological research in a particular elderly setting, however, it was mostly inadequate as a standard two-dose ginkgo efficacy trial in dementia.

Conclusion

In conclusion, the Maastricht study should be considered a pilot trial of ginkgo in a heterogeneous population, at most representative of those elderly with cognitive disorders not otherwise specified, as defined in DSM-IV, living in sheltered centers. It was designed at a past time when AAMI was considered a group of interest for its potential conversion to dementia. Since then, AAMI has been split in very different entities, some of them yet to be validated for their clinical relevance. Furthermore, the last analysis published [5] demonstrated that severity of cognitive impairment, as measured by MMSE, and impairment of social behavior at home, as assessed by caregivers, are important factors to be considered when research aims to evaluate ginkgo effects in patient populations. The particular setting of the Maastricht trial, its unconventional design, and its questionable power computation as well as outcome measurement selection dramatically limit the generalization of its results and unfortunately preclude any comparison with the effect of ginkgo in outpatient population with AD or VaD, as previously studied by Kanowski et al. [3] and Le Bars et al. [4].

References

- 1 Van Dongen MCJM: Efficacy of Ginkgo biloba in dementia and cognitive decline. Ph.D. thesis. Maastricht University, Department of Epidemiology, 1999.
- 2 Van Dongen M, Van Rossum E, Kessels A, Sielhorst H, Knipschild P: The efficacy of Ginkgo for elderly people with dementia and age-associated memory impairment: New results of a randomized clinical trial. *J Am Geriatr Soc* 2000;48: 1183–1194.
- 3 Kanowski S, Herrmann WM, Stephan K, Wierich W, Hörr R. Proof of efficacy of the Ginkgo biloba special extract EGb 761 in outpatients suffering from mild to moderate primary degenerative dementia of the Alzheimer type or multi-infarct dementia. *Pharmacopsychiatry* 1996;29:47–56.
- 4 Le Bars PL, Katz MM, Berman N, Itil TM, Freedman AM, Schatzberg AF. A placebo-controlled, double-blind, randomized trial of an extract of Ginkgo biloba for dementia. *JAMA* 1997;278:1327–1332.
- 5 Le Bars PL, Velasco FM, Ferguson JM, Dessain EC, Kieser M, Hoerr R: Influence of severity of cognitive impairment on the effect of the Ginkgo biloba extract EGb 761 in Alzheimer's disease. *Neuropsychobiology* 2002;45:19–26.