

Efficacy of Ginkgo biloba in dementia and cognitive decline

Citation for published version (APA):

van Dongen, M. C. J. M. (1999). *Efficacy of Ginkgo biloba in dementia and cognitive decline*. [Doctoral Thesis, Maastricht University]. Universiteit Maastricht. <https://doi.org/10.26481/dis.19990326md>

Document status and date:

Published: 01/01/1999

DOI:

[10.26481/dis.19990326md](https://doi.org/10.26481/dis.19990326md)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Efficacy of Gestalt therapy
in scientific and cognitive domains

© M.C.J.M. van Dongen, Maastricht 1999

ISBN 90-90124-68-3

Productie-ondersteuning
Cobie Martens

Vormgeving en druk
Datawyse | Universitaire Pers Maastricht

Promotor
Prof. dr. P.G. Knipschild
Beoordelingscommissie
Prof. dr. M.A. van den Hout (voorzitter)
Prof. dr. J.M.P. Klijn (lid)
Prof. dr. J.A. Kroes (lid)
Prof. dr. J. W.M.M. Bors (lid)
Dr. P.R.

Efficacy of *Ginkgo biloba* in dementia and cognitive decline

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Maastricht
op gezag van de Rector Magnificus,
Prof. dr. A.C. Nieuwenhuijzen Kruseman
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op vrijdag 26 maart 1999 om 14.00 uur

door

Martinus Cornelis Johannes Maria van Dongen

CONTENTS

1	Introduction	9
2	Efficacy of <i>Ginkgo biloba</i> extracts in patients with cerebral insufficiency: evidence from randomized clinical trials	19
3	Efficacy of <i>Ginkgo biloba</i> in elderly people with dementia or age-associated memory impairment: design of a randomized clinical trial	81
4	Ginkgo for elderly people with dementia and age-associated memory impairment: results of a randomized controlled trial	117
5	The efficacy of Ginkgo for elderly people with dementia and age-associated memory impairment: new results of a clinical experiment	133
6	General discussion	155
	Summary	173
	Samenvatting	178
	Dankwoord	188
	Curriculum Vitae	192

1 INTRODUCTION

1 BACKGROUND

In 1993 a report was published that summarized the clinical efficacy of a multitude of therapeutic treatments within the domain of complementary medicine¹. This report was commissioned by the Dutch Ministry of Welfare, Health and Cultural Affairs, and contained a series of reviews dealing with, e.g., the efficacy of EDTA chelation therapy in patients with coronary heart disease, the efficacy of acupuncture as a therapy for various indications, the efficacy of plant products and food supplements such as evening primrose oil, garlic supplements, ginseng, *Ginkgo biloba*, mistletoe and megadoses of the vitamins C, E and B₆, and the efficacy of homeopathy in several disorders. Most of these reviews had also been published as separate articles in the medical literature. All were based on a thorough search of the literature for relevant (randomized) clinical trials, followed by a critical assessment of the methodological quality and the outcomes of each trial. A review of these systematic reviews revealed that the health claims for most of the therapies were supported by a disappointing amount of scientific evidence. One obvious exception, however, caught the eye: *Ginkgo biloba*. The overview that addressed the documented effect of *Ginkgo bi-*

loba special extract in patients suffering from any form of 'cerebral insufficiency' - e.g., dementia, memory impairment, symptoms such as tinnitus, vertigo, headache and depressive mood - included 40 controlled trials, 39 of which reported a beneficial effect of the study medication². And a complementary review showed very consistent, positive results of *Ginkgo biloba* in patients with peripheral arterial occlusive disease as well: all 15 trials included in the review indicated a beneficial effect^{3,4}.

This manifestation of the remarkable position of *Ginkgo biloba* amidst a whole range of other therapies that all had been subjected to a similar, scrutinous assessment, would be the main impetus that eventually made us decide to plan and conduct a new trial in this field: the Maastricht Ginkgo Trial. Several considerations seemed to justify the initiation of yet another trial on the efficacy of *Ginkgo biloba* in patients with cerebral insufficiency:

1. The findings of the Ginkgo review were so impressive, that we almost felt them sound too good to be true.
2. Except for promising results, the review also showed that many of the Ginkgo trials published had suffered from serious methodological shortcomings; the avail-

- able evidence seemed to be based on a rather scarce amount of scientific rigour.
3. In connection with the call for additional studies with an improved methodological quality, it was recognized that the somewhat vague and diffuse concept of cerebral insufficiency should be abandoned. Instead, trials of *Ginkgo biloba* should preferably address rather well-defined and generally accepted diagnostic entities, such as dementia. Moreover, these trials should fulfil the latest requirements with regard to both diagnostic screening procedures and outcome evaluation.
 4. Assuming a beneficial effect of Ginkgo, several additional research questions that have never been seriously addressed by previous trials come into sight. Examples relate to the dose-dependency and the persistency of any effect after a course of *Ginkgo biloba*.
 5. Due to their popularity in several European countries, e.g., Germany and France, Ginkgo preparations represent a high financial value both for manufacturers and consumers. This fact alone warrants additional efforts to prove its clinical efficacy.
 6. The lack of familiarity with Ginkgo products among Dutch health care providers and consumers provides for a favourable setting for a new Ginkgo trial, from a methodological, practical, and ethical point of view.

Quite unanimous notifications with regard to the safety and tolerability of Ginkgo facilitated our decision to commence a new Ginkgo trial. The actual planning and conduct of such a trial became feasible when we obtained financial support from one of the leading manufacturers of *Ginkgo biloba* special extracts. A description of the design and the results of the Maastricht Ginkgo Trial constitutes the backbone of this thesis, together with an updated review of the findings

from previous Ginkgo trials that focussed on cerebral insufficiency, dementia and related disorders.

2 A SHORT HISTORY OF THE USE OF *GINKGO BILOBA* AS A PHYTOMEDICINE

Since its introduction on the market in the mid 1960s, the consumption of *Ginkgo biloba* extract has become increasingly popular in several West-European countries and, more recently, in the USA as well. Positive clinical results of 'Ginkgo' have been claimed for patients suffering from a wide variety of health problems, such as intermittent claudication, Raynaud's syndrome, vertigo, tinnitus, cochlear deafness, senile macular degeneration, diabetic retinopathy, erectile dysfunction, cognitive impairment and dementia⁵⁻⁸. It is a quite heterogeneous list of medical conditions for which Ginkgo preparations are indicated and actually used, but most of them are thought to have a common, vascular origin: an impaired blood supply of the affected organ or tissue.

The first records of the phytotherapeutical use of *Ginkgo biloba* extract go back to ancient China. Approximately 3000 BC, Ginkgo was listed in the first known pharmacopoeia, together with 365 other medical substances⁹. Ginkgo leaves and nuts were recommended for combatting disorders such as asthma, bronchitis, micturition problems and spermatorrhoea. Applications of *Ginkgo biloba* extracts, most often administered as a herbal tea, have been reported from several countries in East-Asia (China, Japan, Korea) throughout the Middle Ages. These extracts were used for various medical conditions, apart from asthma and bronchitis, also tuberculosis, gonorrhoea, stomach complaints, skin diseases, etc. The prominent role of *Ginkgo biloba* in popular medicine was fully

in agreement with the 'sacred' status of the tree. Throughout the centuries myth and reality went hand in hand to emphasize the unique character and strength of the tree and its products. *Ginkgo biloba* is believed to be one of the oldest, still existing plants. Specimens have been identified that have survived for more than 1,000 years. The story goes that a Ginkgo tree, after having been burnt during the atom bomb assault on Hiroshima in 1944, produced a new shoot already a year later. Moreover, the Ginkgo tree is supposed to be highly resistant against all kinds of phytopathological disasters, whether they are caused by moulds or by bacterial or viral infections. Around 1730 the Ginkgo tree was reintroduced in Europe - in Holland - where it had disappeared during the Ice Age. However, it was cultivated here just for ornamental purposes, at least for the time being. Its medicinal properties would be recognized again only in the 1960s, after the new discoveries and developments in analytical chemistry (e.g., chromatography, spectroscopy) had enabled the identification and isolation of distinct active compounds stored in the various parts of the Ginkgo tree.

In 1965 the first standardized Ginkgo extract was marketed in Europe, by Dr. Willmar Schwabe GmbH at Karlsruhe, Germany, under the brand name Tebonin®. Soon, it was followed by other products, e.g., Tanakan® (IHB/Ipsen, Paris), which is mainly distributed on the French market, rökan® (Intersan, Ettlingen), Tebonin® forte (Schwabe, Karlsruhe), and Kaveri® (Lichtwer Pharma, Berlin). The production process that yields the active compound for these preparations is aimed at the extraction, concentration and standardization of two principal groups of chemical substances: flavonoids and terpenes. Flavonoids are very common in the vegetable kingdom. The Ginkgo tree contains some characteristic flavone- and flavonol glycosides - in general with quercetin,

kaempferol, and isorhamnetin as the aglyconic part - as well as biflavonoids, e.g., isoginkgetin¹⁰. The Ginkgo flavonol glycosides have been postulated to act as scavengers of free radicals. These free radicals may cause a reduced blood circulation through a combination of mechanisms: peroxidation of membrane lipids, increased aggregation of thrombocytes, edema leading to metabolic disturbances, and vessel wall damage. The terpene fraction of *Ginkgo biloba* extract contains some unique diterpenes (Ginkgolides A, B, C and J) and the sesquiterpene bilobalide. These constituents, with Ginkgolide B probably being the most specific and active one, have proved to act as strong antagonists of platelet activating factor (PAF), a phospholipid that provokes several physiological reactions through a number of subsequent steps, e.g., aggregation of thrombocytes, vasodilation, bronchoconstriction, inflammation and various anaphylactic reactions^{9,11}.

3 PHARMACOLOGICAL MECHANISMS

Detailed reviews of the pharmacological actions of *Ginkgo biloba* special extracts which might be responsible for the clinical effects evaluated in this thesis, have been presented by other authors^{7,11-19}. In summary, it is assumed that *Ginkgo biloba* constituents may:

- Evoke an increased tolerance to hypoxia, especially in brain tissue.
- Act as free radical scavengers, resulting in the prevention of membrane damage, e.g., damage to the erythrocyte membranes (scavenger of toxic oxygen derivatives, such as superoxide, hydroxyl, peroxy, lipid peroxidation products).
- Prevent the development of brain edema (post-traumatic, toxin-induced).

- Improve the rheological properties of the blood and exert vasoregulating activity.
- Act as a platelet activating factor (PAF) inhibitor and exert a neuroprotective effect, under conditions such as hypoxia / ischemia, seizure activity and peripheral nerve damage. Either direct influences on CNS neurons or indirect effects on the CNS via modulation of the blood flow may be responsible for this protective action. Both the terpene fraction of *Ginkgo biloba* extract, containing Ginkgolide B, and the flavonoid fraction may contribute to the neuroprotective action.
- Inhibit the age-related decline of muscarinic choline receptors and promote the choline uptake in the hippocampus.
- Reduce retinal edema.

These effects may be evoked either by one or more single active ingredients, or by the concerted action of several of the many potentially active agents that can be detected in *Ginkgo biloba* constituents.

4 GINKGO BILOBA SPECIAL EXTRACTS

Initially, medical preparations based on *Ginkgo biloba* extracts were standardized exclusively on their flavonol glycoside content. Comparable standards for the terpene content were added later. Nowadays, the most popular Ginkgo preparations contain 24 mg Ginkgo flavonol glycosides and 6 mg terpene trilactones per 100 mg extract. These concentrations are obtained by a process of enrichment and purification of the original, full extract from dried, fragmented Ginkgo leaves, that normally contains 2-4% of Ginkgo flavonol glycosides (standard: at least 2%). Extracts that have been subjected to this process of enrichment are called *Ginkgo biloba* special extracts (GBSE). The flavonol glycoside content of these special

extracts varies between 16% and 25%²⁰. EGb 761 (Extractum Ginkgo bilobae 761) is used as a code name to identify the most familiar special extract, which is standardized on 24% Ginkgo flavonol glycosides and 6% terpene trilactones. The registered drugs Tebonin[®], rökän[®] and Tanakan[®] are all based on EGb 761. LI 1370 stands for another special extract, standardized on 25% flavonol glycosides and 6% terpene trilactones, and with the trademark Kaveri[®] connected to it. The great majority of the intervention trials which have been conducted until now to assess the efficacy of Ginkgo treatment in various health domains, have used either EGb 761 or LI 1370 as the experimental agent. Nowadays *Ginkgo biloba* special extracts, often EGb 761, are also part of an increasing number of newer products, marketed by pharmaceutical companies all over the world, e.g., Ginkgold[™], Bio-Biloba[™] (Pharma Nord, Denmark), Ginkoba[™] (Pharmaton, USA), Tebofortan[®] (Austria), and many more.

Ginkgo biloba extract can be administered in various ways. Usually it is supplied per os, either in a solid form (film-coated tablets, dragees, capsules), or as a solution (drops). Moreover, in clinical settings intravenous or intramuscular injections, and intravenous infusions are alternative options, which have been preferred in several intervention trials. Tablets and solutions may contain various amounts of the extract that harbours the active compounds. Eventually, the expected effect will depend on the total daily dose and the time-related distribution of this dose (number and timing of intakes).

For example, regular dosage schedules for various Tebonin[®] modalities are: film-coated tablets: 3 tablets a day (1 tablet = 40 mg of active substance (dry extract; input : output ratio = 50 : 1)); drops: 3 x 20 drops a day (20 drops = 1 ml of solution = 40 mg of active substance); intravenous or intramuscular injection: 1 or 2 times 1 flask a day (1 flask = 3

ml ready-for-use solution = 50 mg of active substance); intravenous infusion: 2-4 flasks a day, either as a single dose or distributed over the day (1 flask = 3 ml ready-for-use solution = 50 mg of active substance).

In the past a total daily amount of 80-160 mg special extract, usually corresponding with 1-4 tablets, was regarded to be sufficient, but recently a tendency can be noticed to elevate the dose a little bit. The discussion on the optimal dose is facilitated by the fact that *Ginkgo biloba* preparations are considered to be very safe and highly tolerable, a supposition that has been confirmed time and again in clinical studies and through post-marketing surveillance.

The safety and tolerability of *Ginkgo biloba* extract, combined with the claims for its health enhancing capacities, seem to favour the incorporation of Ginkgo constituents in multi-preparations and all kinds of health products that promise to fortify the inner or outer man. And this is exactly what has happened during the past decade. Nowadays Ginkgo compounds act as ingredients of a wealth of preparations that can be obtained without any restriction - except for the financial part of the deal - not as drugs that are prescribed by the doctor, but as over-the-counter medicines, via pharmacies, reform houses, health food stores, and other canals. Often, these Ginkgo compounds are not derived from one of the special extracts just mentioned, but from the full extract, without subsequent concentration. As a consequence, consumption of these products according to the labelled instructions will normally result in a daily intake that lags far behind the dose that is claimed to be effective by experts. In 1995 an investigation conducted by the Dutch Consumers Association²¹, in which 13 Ginkgo preparations available on the Dutch market were involved, revealed that most of them - often promoted for people with geriatric complaints (e.g., Geriaforce, Geria+, Geria-totaal) - did not

meet at all the generally accepted standards regarding the daily dose of the active substances. This development asks for special efforts to control the quality of Ginkgo preparations that appear on the market²².

5 USE, PRESCRIPTIONS AND SALES FIGURES OF GINKGO PREPARATIONS

The use of phytomedicines has shown a drop in many Western countries during the last three decades, with the arrival of newly developed synthetic chemicals and the efforts to synthesize structural analogs of plant products²³. However, there are signs that the 'botanical dark age' is beginning to lift also in the USA, both because of the potentials of the rational use of natural products and the limitations with regard to efficacy and safety of synthesized medicines, especially for psychotropic indications. Some observers consider Western Europe the trendsetter for future developments in this area²³. Their remarks apply to Germany in particular, where phytomedicine is a regular part of mainstream medicine, often reimbursed by the health care system. Botanical medicine receives an ever increasing amount of attention, not only from individual consumers and consumers' organizations, but also from the health authorities. Recently, this was illustrated by a joint meeting of the Food and Drug Administration, the National Institutes of Health Office of Alternative Medicine and the herbal industry, which was organized in December 1994, in Washington, D.C. At this meeting agreements were made on the responsible and self-regulating role of the industry, the process of quality assurance of herbal products, the establishment of quality guidelines, and the continuous reviewing and assessment of efficacy and safety. An editorial comment in the *Journal of the American Medical Association*, which sum-

marized the meeting, gave its support to the German system for evaluating the efficacy and safety of botanical medicines, with the Commission E as a regulatory authority that deals exclusively with botanicals^{23,24}. The actual role that phytomedicines play on the market of health products and the share of Ginkgo preparations can be illustrated with several figures. In Germany Ginkgo preparations belong to the most frequently prescribed drugs already for a long time, with, for instance, more than 5 million prescriptions in 1988, at a cost for the health insurance of 370 million DM^{25,26}. In 1989, the cost for Tebonin alone was 219 million DM, and for rōkan 62 million DM, while the costs for the antihypertensive agent Adalat (nifedipine), number two on the ranking, amounted to 195 million DM. Figures presented at the US-meeting in 1994, revealed that the total phytomedicine sales - in millions US \$ - in 1992 were estimated to be 2,300 for the Asian region, 2,100 for Japan, 1,500 for Germany (more than 50% prescribed by physicians), 1,000 for North America, 210 for France, 190 for the UK, and 92 for the Netherlands^{23,27}. The leading European product was Tebonin (Schwabe) with a sales figure of 195 million US \$, followed by the Panax ginseng product Ginsana (Pharmaton / Boehringer Ingelheim; 50 million US \$), and the garlic product Kwai (Lichtwer Pharma; 40 million US \$). The approval of Tebonin as an anti-dementia drug by the German health authorities in 1994, was expected to have a strong upward influence on the sales figures. In 1994 *Ginkgo biloba* was again with a large margin the most prescribed herbal monopreparation in Germany, with more than 4 million prescriptions at a total price of 263 million DM²⁷. A few years ago more than a dozen *Ginkgo biloba* products were spotted on the USA market²⁸. Among them were several special extracts, such as Ginkgo Power™, Super Ginkgo™, Ginkgold™ and Ginkoba™ (Phar-

maton). The fact that herbal medicine is a booming business and that the historical interest by pharmaceutical companies has changed is also illustrated by the fact that recently many relatively small botanical and herbal medicine companies have been taken over by multinational pharmaceuticals^{23,27}.

Ginkgo biloba special extracts have officially been registered as drugs in several European countries, such as Germany, France, and Austria. The status which *Ginkgo biloba* preparations and other registered nootropics have acquired as cognition enhancers (nootropic drugs) in Germany, is nicely illustrated by a survey conducted in Lower Saxony^{29,30}. A sample of general practitioners and neuropsychiatrists was randomly confronted with two written case vignettes, one representing a case with mild memory impairment, either with or without the expressed wish for medication, and the other representing a case of dementia, either with a typical history of vascular dementia or with a typical history of Alzheimer's dementia. Regardless of the wish of the patient and the type of decreased brain function, 70% of the participating physicians declared that they would prescribe any of the given nootropic drugs, although 58% expressed doubts about their effectiveness, and although only 28% indicated to expect a positive effect on brain performance. Among the family physicians - but not among the neuropsychiatrists - *Ginkgo biloba* turned out to be the favourite drug for treating the dementia case, irrespective of the type of dementia (24% would prescribe Ginkgo for case A, and 28% for case B)^{29,30}.

6 EFFICACY OF GINKGO IN CEREBRAL INSUFFICIENCY: EVIDENCE FROM PREVIOUS TRIALS

As we have stated before, quite a large number of controlled trials have been carried out during the past two decades to elucidate the efficacy of the *Ginkgo biloba* preparations. We mentioned already the systematic review of Kleijnen and Knipschild, that summarized the results and assessed the methodological quality of 40 randomized clinical trials on *Ginkgo biloba* and cerebral insufficiency^{2,4}. This review showed that the intervention period lasted 6-12 weeks in most of the trials, with a daily dose of 120-160 mg (frequency of dosage: two to three times a day). The methodological quality of the trials was found to vary considerably: total scores awarded on a rating scale with a minimum value of 0 points and a maximum value of 100 points, ranged from 18 till 90. Several shortcomings were noticed repeatedly: inadequate sample size; lack of control over the maintenance of the blinding procedure; insufficient description of randomization procedures, characteristics of the study population and effect measures; and an incomplete presentation of the study results. Apart from a single exception all trials tended to reveal a positive effect of Ginkgo. In many trials the measured responses appeared large enough to be of clinical relevance. Moreover, the results seemed to match those reported in a series of methodologically comparable trials which studied the pharmacotherapeutical effects of codergocrine (Hydergin[®]; a registered drug in the Netherlands). Unwanted side-effects of *Ginkgo biloba* extract were not reported in any of these studies. The experiences available from these previous trials made up the scenery for the design of the Maastricht Ginkgo Trial.

7 STRUCTURE OF THE THESIS

The remainder of this thesis is structured as follows.

In chapter 2 we review the currently available evidence for the efficacy of *Ginkgo biloba* special extracts in patients suffering from various types of cerebral insufficiency, including dementia and cognitive decline. Although we have chosen a different strategy to assess and summarize the evidence from published trials on Ginkgo and cerebral insufficiency, our review can be considered an update of previous literature reviews dealing with the same subject. Special attention is paid to the results of the latest trials, which primarily focussed on the treatment of dementia. Chapter 3 describes the design of a new randomized, placebo-controlled, double-blind, clinical trial in Maastricht, the Netherlands. This trial was initiated in 1993 to study the efficacy and the safety of the *Ginkgo biloba* special extract EGb 761 in elderly people suffering from mild to moderate dementia or age-associated memory impairment.

In chapter 4 we present the main results of the Maastricht Ginkgo Trial. These results refer to the primary outcome measures that we used to evaluate the efficacy of Ginkgo. Additional effects of Ginkgo, based on a range of supplementary outcome measures, are addressed in chapter 5.

Chapter 6 contains a general discussion on the evidence that we have been able to collect regarding the efficacy of *Ginkgo biloba*. Firstly, we critically reflect on the information contained by our systematic review of all previous Ginkgo trials. Subsequently, we assess the significance of the results of our own trial, in view of its validity and precision. Thereafter, we attempt to reconcile the findings from the Maastricht Ginkgo trial with those from previous trials. Finally, I will express my personal view regarding the efficacy of *Ginkgo biloba* in dementia and other manifestations of cerebral insufficiency.

REFERENCES

1. Kleijnen J, Ter Riet G, Knipschild P, eds. *Effectiviteit van alternatieve geneeswijzen; een literatuuronderzoek*. Maastricht: Rijksuniversiteit Limburg, Vakgroep Epidemiologie, 1993.
2. Kleijnen J, Knipschild P. Ginkgo biloba for cerebral insufficiency. *Br J Clin Pharmacol* 1992; 34: 352-8.
3. Kleijnen J, Knipschild P. Ginkgo biloba for intermittent claudication and cerebral insufficiency. In: Kleijnen J, ed. *Food supplements and their efficacy*. Maastricht: Rijksuniversiteit Limburg, 1991: 83-94.
4. Kleijnen J, Knipschild P. Ginkgo biloba. *Lancet* 1992; 340: 1136-9.
5. Füntgeld EW, ed. *Rökan (Ginkgo biloba). Recent results in pharmacology and clinic*. New York: Springer-Verlag, 1988.
6. DeFeudis FV. *Ginkgo biloba extract (EGb 761). Pharmacological activities and clinical applications*. New York: Elsevier, 1991.
7. DeFeudis FV. *Ginkgo biloba extract (EGb 761). From chemistry to the clinic*. Wiesbaden: Ullstein Medical, 1998.
8. De Munck-Khoe LK. Ginkgo biloba. Heilzame erfenis uit de oertijd. *Nieuwsbrief Stichting Science & Nutrition* 1997: 1-14.
9. Z'Brun A. Ginkgo - Dichtung und Wahrheit. *Schweiz Rundsch Med Prac* 1995; 84: 1-6.
10. Hölzl J. Inhaltsstoffe von Ginkgo biloba. *Pharm Unserer Zeit* 1992; 21: 215-23.
11. Oberpichler Schwenk H, Krieglstein J. Pharmakologische Wirkungen von Ginkgo biloba Extrakt und Inhaltsstoffen. *Pharm Unserer Zeit* 1992; 21: 224-35.
12. Hitzenberger G. Zur Wirkung von Ginkgo biloba Spezialextrakt (EGb 761, Tebofortan). *Wien Med Wochenschr* 1992; 142: 371-9.
13. Ferradini C, Droy-Lefaix M, Christen Y, eds. *Ginkgo biloba Extract (EGb 761) as a free-radical scavenger*. Paris: Elsevier, 1993.
14. Clostre F, DeFeudis F, eds. *Cardiovascular effects of Ginkgo biloba Extract (EGb 761)*. Paris: Elsevier, 1994.
15. Christen Y, Droy-Lefaix M, Macias-Nunez J, eds. *Effects of Ginkgo biloba Extract (EGb 761) on neuronal plasticity*. Paris: Elsevier, 1996.
16. Christen Y, Costentin J, Lacour M, eds. *Effects of Ginkgo biloba Extract (EGb 761) on the central nervous system*. Paris: Elsevier, 1992.
17. Anonymous. Tebonin® (EGb 761). *Ginkgo biloba Special Extract. The result of 25 years of research*. Karlsruhe: Dr. Willmar Schwabe GmbH & Co, 1991.
18. Packer L, Christen Y, eds. *Ginkgo biloba extract (EGb 761): lessons from cell biology*. Paris: Elsevier, 1998.
19. Papadopoulos V, Drieu K, Christen Y, eds. *Adaptive effects of Ginkgo biloba extract (EGb 761)*. Paris: Elsevier, 1997.
20. Sticher O. Ginkgo biloba Analytik und Zubereitungsformen. *Pharm Unserer Zeit* 1992; 21: 253-65.
21. Anonymous. Meeste echinacea- en Ginkgo-middelen twijfelachtig. *Consumentengids* 1995: 528-9.
22. Sticher O. Quality of Ginkgo preparations. *Planta Med* 1993; 59: 2-11.
23. Cott J. NCDEU update. Natural product formulations available in Europe for psychotropic indications. *Psychopharmacol Bull* 1995; 31: 745-51.
24. Marwick C. Growing use of medicinal botanicals forces assessment by drug regulators. *JAMA* 1995; 273: 607-9.
25. Anonymous. Ginkgo biloba extract: over 5 million prescriptions a year. *Lancet* 1989; ii: 1513-4.
26. Schwabe U, Paffrath D. *Arzneiverordnungs-Report '90*. Stuttgart: Gustav Fisher, 1989.
27. Grünwald J. The European phytomedicines: market figures. *Herbalgram* 1995; 34: 60-5.
28. Itil T, Martorano D. Natural substances in psychiatry (Ginkgo biloba in dementia). *Psychopharmacol Bull* 1995; 31: 147-58.
29. Stoppe G, Sandholzer H, Staedt J, et al. Einflußfaktoren auf die Verordnung von Nootropika. Ergebnisse einer repräsentativen Umfrage in Niedersachsen. *Dtsch Med Wochenschr* 1995; 120: 1614-9.
30. Stoppe G, Sandholzer H, Staedt J, Winter S, Kiefer J, Ruther E. Prescribing practice with cog-

nition enhancers in outpatient care: are there differences regarding type of dementia? - Results of a representative survey in Lower Saxony, Germany. *Pharmacopsychiatry* 1996; 29: 150-5.

Accuracy of Ginkgo biloba extracts in patients with cerebral insufficiency: evidence from randomized clinical trials

© 1997 Taylor & Francis Ltd.

2 Efficacy of *Ginkgo biloba* extracts in patients with cerebral insufficiency: evidence from randomized clinical trials

Martien van Dongen, MSc¹

Erik van Rossum, PhD¹

Paul Knipschild, MD, PhD²

¹ Department of Epidemiology, Maastricht University

² Department of General Practice, Maastricht University

This chapter is an adapted version of: Van Dongen M, Van Rossum E, Knipschild P. Efficacy of Ginkgo biloba extracts: evidence from randomized clinical trials. In: Van Beek TA, ed. Ginkgo biloba. Amsterdam: Harwood Academic Publishers, 1999 (In press).

ABSTRACT

Objective

This review summarizes the available evidence for the clinical efficacy of *Ginkgo biloba* extract in patients with various forms of cerebral insufficiency, including dementia.

Method

By means of a thorough literature search we have attempted to identify all published randomized intervention trials which assessed the effect of *Ginkgo biloba* special extract on clinical outcome parameters in human patients with manifestations of cerebral insufficiency. For this purpose, we performed computer aided searches of bibliographic databases, we went through conference proceedings and abstracts and through the contents of relevant journals, we checked references in publications on Ginkgo, and we consulted Ginkgo experts. Several considerations made us to abstain from both a systematic review by means of criteria-based methodological assessment of each eligible study, and a quantitative meta-analysis based on statistical pooling of individual study data. Instead, we preferred to group the identified trials together into a number of more or less homogeneous categories according to the health problem addressed, and to present a detailed description of the main design characteristics and results of each study, accompanied by a summary judgment of the trials within each category.

Results

We found 55 randomized clinical trials that investigated the effects of *Ginkgo biloba* special extract in patients with cerebral insufficiency. According to the target health problem addressed the following headings were used to categorize these trials: dementia (9 trials), serious cognitive impairment (8 trials), non-cognitive signs and symptoms of cerebral insufficiency (12 trials), subjective cognitive and non-cognitive symptoms of cerebral insufficiency (5 trials), depression (3 trials), and cerebral insufficiency unspecified (18 trials). Beneficial effects of Ginkgo were reported for almost all trials in each category, although not invariably for all outcome measures. The strongest effects were reported for the older trials, whereas the newer trials tend to be superior as to methodological quality.

Conclusion

The available trials suggest that *Ginkgo biloba* special extract is efficacious in patients with dementia and other forms of cerebral insufficiency. However, the clinical meaningfulness of the reported effect sizes remains difficult to interpret. Since only a limited number of trials fulfils the current standards for evaluation and reporting, there is still need for additional information from rigorously designed clinical trials.

1 INTRODUCTION

The aim of this chapter is to review the evidence regarding the clinical efficacy of *Ginkgo biloba* extract that is currently available in the scientific literature.

We have decided to restrict the scope of the review in several ways. Firstly, it will deal exclusively with intervention studies in human patients with manifestations of cerebral insufficiency. It will not cover the evidence gathered through animal studies or in vitro investigations. Studies involving healthy volunteers will not be taken into consideration either. Only those health problems will be covered that are considered consequences of an insufficient cerebral blood flow or any other lesion in the brain area (e.g., metabolic disturbances, cell membrane damage). Dementia (Alzheimer's disease, vascular dementia), cognitive impairment being no dementia (age-associated memory impairment (AAMI), age-associated cognitive decline (AACD)), and several non-cognitive sequelae of an inadequate blood supply of the brain region - such as vertigo, tinnitus, hearing loss, and headache - are the main disease phenomena addressed in the current review. An overview of the evidence of the efficacy of *Ginkgo biloba* in patients suffering from other circulatory disorders, in particular peripheral arterial obstructive disease (PAOD), will be presented elsewhere¹.

Secondly, only those studies will be addressed that focus predominantly on 'clinical' intervention endpoints, such as general well-being, clinical signs and symptoms, and functional capacity. Studies with an exclusive focus on metabolic or physiological effects of *Ginkgo biloba* products and constituents will be omitted from the review, as they generally aim at the disclosure of the mechanisms of action underlying the clinical effects of Ginkgo treatment, not of the clinical efficacy as such. Detailed reviews of the pharmacological actions of *Ginkgo biloba* special

extracts can be found elsewhere²⁻⁶. In summary, it is assumed that Ginkgo constituents may:

- improve the rheological properties of the blood and exert a vasoregulating activity, giving rise to an increased blood flow;
- act as a platelet activating factor (PAF) inhibitor and exert a neuroprotective effect, under conditions such as hypoxia / ischemia, seizure activity, and peripheral nerve damage; either direct influences on CNS neurons or indirect influences on the CNS via modulation of the blood flow might be responsible for this protective action; both the terpene fraction of *Ginkgo biloba* extract, containing Ginkgolide B - a very potent PAF antagonist - and the flavonoid fraction might contribute to the neuroprotective action;
- act as a free radical scavenger, preventing membrane damage, e.g., damage to erythrocyte membranes (scavenger of toxic oxygen derivatives, such as superoxide, hydroxyl, peroxy, lipid peroxidation products);
- prevent the development of brain edema (post-traumatic, toxin-induced);
- evoke increased tolerance to hypoxia, especially in the brain tissue;
- inhibit the age-related decline of muscarinic choline receptors and promote the choline uptake in the hippocampus.

Thirdly, the review is concentrating on published evidence from investigations designed to exert a sufficient amount of control over potential sources of bias. This means that only the results of randomized clinical trials will be presented and discussed. Evidence that is based on case series, open trials, and other types of design, is neglected. A final restriction of the present review concerns the type of intervention evaluated: the review includes only studies intended to assess the efficacy of a *Ginkgo biloba* special extract (GBSE), that means an extract con-

taining fixed amounts of the substance classes that are regarded to be responsible for its clinical efficacy (in general: 24-25% Ginkgo flavonol glycosides, 6% terpene trilactones), obtained through a sophisticated and highly standardized process of extraction, purification and concentration⁷.

The review is structured as follows. Firstly, we will briefly reflect on the results of and experiences acquired from previous reviews of *Ginkgo biloba* and cerebral insufficiency. Subsequently, we will explain and legitimize the strategy chosen for the current review. Thereafter, we will present a systematic overview of the intervention studies of *Ginkgo biloba* and cerebral insufficiency that we have been able to retrieve. For this purpose we have grouped the eligible trials into several categories, according to the target health problem addressed. The main characteristics of the trials in each category will be communicated in a tabulated format. Each table will be commented on briefly. Finally, we will attempt to formulate a summary judgment on the efficacy of *Ginkgo biloba* in patients with dementia, cerebral insufficiency and related disorders.

2 PREVIOUS REVIEWS OF *GINKGO BILOBA* AND CEREBRAL INSUFFICIENCY

Several types of literature review can be distinguished in evaluating the efficacy of medical interventions⁸⁻¹⁰.

The classic approach to research synthesis is sometimes called the 'narrative' review. Often such a review is composed by a subject-matter expert. It bears the stamp of the authority and the personal preferences of the reviewer, and often lacks explicit statements regarding the searching strategy applied, the inclusion and exclusion criteria used to select studies, and the main argu-

ments which gave rise to the summary judgment. In order to circumvent the limitations inherent to this classic approach of reviewing, more systematic and explicit approaches have been developed during the past two decades. Two modern types of systematic literature review that have become popular are the criteria-based systematic review and the quantitative meta-analysis.

The criteria-based systematic review - or qualitative meta-analysis - stresses the need for a critical and structured assessment of the methodological quality of each study that qualifies for the review. Sometimes this assessment is combined with a scoring system that is suited to value each methodological aspect separately and to calculate a summary quality score for the study at large. The idea behind this approach is that the methodological quality should play a major role in weighing the evidence of each of the studies, in order to prevent a biased summary judgment.

The quantitative meta-analysis honours the principle of statistical pooling of the results of several intervention trials addressing the same research question. Statistical pooling allows for an overall estimation of the size of the intervention effect which is more precise than the effect estimations for each of the studies separately. Statistical pooling of individual study outcome data and systematic, criteria-based methodological quality assessment should not be considered mutually exclusive strategies. On the contrary, they can be combined and integrated in a fruitful way to enhance both the precision and the validity of the final summary judgment.

Efforts to extract a valid summary judgment from the literature may be hampered in several ways. First of all, selective publication of empirical evidence and selective retrieval of published evidence may result in biased conclusions, irrespective of the type of systematic review. Heterogeneity of the methodological quality of a series of studies ad-

addressing the same research question, is a second obstacle. This problem is best addressed by means of a criteria-based systematic review. Heterogeneity of studies with regard to various design aspects that not necessarily affect the methodological quality - e.g., the choices of study subjects, type and duration of the intervention, and outcome measures - can be regarded as a third sort of problem, which especially interferes with statistical pooling.

Several reviews assessing the efficacy of *Ginkgo biloba* in patients with cerebral insufficiency and related disorders have been published in the recent past. We are inclined to classify most of them under the heading 'narrative review'¹¹⁻¹⁸, although some tend to a systematic approach with a strong narrative component. Two clear exceptions are a criteria-based systematic review¹⁹⁻²¹, and a meta-analysis based on the pooled data of 8 randomized placebo-controlled studies²². For the methodological appraisal of the Ginkgo trials in patients with cerebral insufficiency that could be retrieved by the end of 1991, Kleijnen et al applied a 100-point assessment scale covering seven methodological criteria. Details regarding the scoring system can be found in the original publications^{19,20}. Of the 40 studies included in the review, only 8 were classified as 'acceptable' in terms of their methodological quality (total score ≥ 65)²³⁻³⁰. No less than 39 studies indicated a positive effect of Ginkgo, which was specified as statistically significant for 26 studies, and as a positive trend for 13 studies. Therefore, study outcome appeared not to be dependent on methodological quality. An insufficient sample size, an inadequately described randomization procedure, and a poor description of patient characteristics and outcome measures were identified as the major methodological shortcomings. Statistical pooling of the results of Ginkgo trials in patients suffering from cerebral insuffi-

ciency is discouraged by their remarkable heterogeneity with regard to relevant design aspects. This heterogeneity concerns in particular the patient definitions used (various diagnostic screening tools, various criteria for inclusion and exclusion) and the outcome measures applied (different levels of observation; various rating scales, psychometric tests and other outcome measures). The intervention procedures and contrasts tend to show more similarity. Whereas the criteria-based systematic review by Kleijnen and Knipschild²⁰ in 1992 contained 40 'cerebral insufficiency' trials, only 11 were considered eligible for the quantitative meta-analysis published by Hopfenmüller²² in 1994, which concentrated on the *Ginkgo biloba* special extract LI 1370 (Kaveri[®] forte) and on a subjective, clinical symptom checklist as the main outcome measure. Eight trials qualified for the analyses based on the evaluation of single symptoms of cerebral insufficiency and the total score for clinical symptoms. Statistically significant effects were found for all symptoms separately and in combination (total score), except for one trial³¹. This indicates the superiority of Ginkgo over placebo. After dichotomization of the total scores ('improved' vs 'not improved') a pooled odds ratio of 2.0 (95% confidence interval: 1.4 ; 2.6) was calculated. Six trials could be used for the meta-analysis based on global assessment of the treatment effect, either by the physician or by the patient, as the main outcome measure. After dichotomization ('very good' or 'good' vs 'moderate' or 'unsatisfactory' or 'no effect') a pooled odds ratio of 2.9 (95% confidence interval: 2.1 ; 3.7) was calculated. These data suggest that the rate of success for Ginkgo treatment is about two to three times that for placebo treatment. Although Hopfenmüller's review illustrates the feasibility of statistical pooling for Ginkgo trials, this could only be attained at the cost of the exclusion of valuable evidence embedded in trials which were not admissible to

the 'pool', due to unconciliatory design choices²².

Actually, all reviews mentioned above agreed upon the beneficial effect of Ginkgo in patients suffering from cerebral insufficiency.

3 DESIGN OF THE CURRENT REVIEW

3.1 REVIEW STRATEGY: GENERAL CONSIDERATIONS

For the current review we decided to apply neither the principles of criteria-based quality assessment, nor the rules of statistical pooling. Instead we preferred to steer more or less a middle course between the narrative approach and the criteria-based quality assessment approach. Through several tables, organized according to the type of health problem of interest, we will present relevant information on the crucial design aspects as well as on the numerical and statistical results for each *Ginkgo biloba* intervention trial that fulfilled the selection criteria mentioned before. Each table will be accompanied by an overall judgement of the meaningfulness of the studies enclosed.

For several trials we found it far from easy to assess to what extent the criteria for inclusion were met, due to a lack of specific information in the publication. We gave such studies the benefit of the doubt and included them in the overview. Our way of summarizing the information is to some extent comparable with the approach applied in two earlier reviews^{12,13}, but we have chosen to present more details in the tables and less in the text, and we have added the most recent trials. The considerations which motivated the current approach will be explained below.

3.1.1 Completeness and transparency of study reports

Many studies published in the past suffer from a rather poor reporting style (non-transparent structure, lack of essential information). This makes it difficult for a reviewer to distinguish between a poor design and conduct of the trial at the one hand, giving rise to invalid results, and a poor description of probably valid results at the other hand. One should not blame the authors too much for this, as the criteria for reporting clinical trials in general, and trials in the field of cognitive functioning and dementia in particular, have evolved gradually during the last few decades³². Nevertheless, this circumstance seriously hampers any attempt to perform an authoritative criteria-based methodological quality assessment.

Provided that sufficient information regarding a clinical trial is forwarded through a well-structured publication that meets the current reporting standards, the methodological assessment of most of the design aspects, such as random allocation, placebo-intervention, and blinded outcome measurement, is rather straightforward. For *Ginkgo biloba* trials the trouble starts with the critical assessment of the patient definition (diagnostic criteria used for the selection of study subjects) and the outcome measurement (evaluation criteria used for identifying the treatment response). These problems are more pronounced for clinical Ginkgo studies focussing on cerebral insufficiency than, for instance, peripheral arterial disease.

3.1.2 Diagnostic uncertainties

Before discussing the methodological problems regarding outcome measurement, we have to elaborate on the diagnostic difficulties and uncertainties surrounding the identification of the target conditions for Ginkgo treatment. Most of the Ginkgo trials in the

past, mainly from Germany and France, were conducted in patients identified with cerebral insufficiency. Cerebral insufficiency and related concepts, such as cerebrovascular insufficiency and brain-organic psychosyndrome (German: HOPS), have been regarded as a separate disease entity for quite a long time. The use of these terms and concepts reflected the contemporary etiological hypothesis that stenotic vascular changes, related to the ageing process, may cause a progressive decrease in cerebral blood flow, leading to a decline in mental and physical functioning. The clinical manifestations of this pathophysiological process were thought to encompass impairment of memory and other cognitive functions, affective symptoms such as anxiety and depression, and physical complaints such as tinnitus, vertigo, and headache^{15,18}.

During recent years this etiological concept of cerebral insufficiency has become obsolete at a fast pace, especially after one came to recognize that neuronal degeneration as in Alzheimer's disease is a more frequently occurring cause of cognitive impairment in elderly patients than vascular deficiency. The clinical features of these central nervous system disorders correspond to the syndrome of dementia, a pattern of disturbance in which several higher mental functions are affected simultaneously, with impairment of memory, abstract thinking, and psychomotor functions like speech as crucial symptoms, and irreversible changes in mood, social functioning, and personality as accompanying phenomena^{15,18}. The criteria for the (differential) diagnosis of dementia, dementia subtypes - Alzheimer's disease (AD) due to an idiopathic degeneration of the central nervous system, vascular dementia (VD; previously called multi-infarction dementia (MID)) due to changes in the vascular system, mixed forms of AD and VD, dementia due to other causes, such as AIDS, brain tumors, et cetera - and related diseases have

evolved steadily during the past 15 years³³⁻³⁸.

To satisfy the need for disease classification within the domain adjacent to dementia and obviously not covered by the diagnostic criteria for dementia, the concept of age-associated memory impairment (AAMI) has been proposed. This concept refers to a pre-dementia state of impairment of memory function and other cognitive functions (language, orientation, constructional abilities, abstract thinking, problem solving, praxis), established both by means of psychometric testing and subjective reporting, and not (yet) accompanied by impairment of social functions (occupational or social performance). Although AAMI and related concepts - such as cognitive impairment no dementia (CIND) and age-associated cognitive decline (AACD) - have always remained controversial as diagnostic entities, formal criteria to identify cases affected by this syndrome have been suggested at least^{39,40}. Eventually, the concept of age-associated cognitive decline has acquired an autonomous state in the DSM-IV classification system³⁶.

Given these nosological developments, it is not surprising that nowadays a more specific and focussed diagnosis is required for those disorders that were earlier classified under the heading 'cerebral insufficiency'. The majority of the Ginkgo trial reports that we could retrieve from the literature do not allow for such a differentiation. It would not be fair to blame the investigators, as most of the trials have been done at a time that the new criteria, guidelines and screening tools for such a differential diagnosis were still awaiting approval and publication. Even nowadays the diagnosis of cognitive neurological disorders is surrounded by much uncertainty and dissent. As a consequence only some of the latest trials are believed to meet the current standards for patient selection in anti-dementia drug trials¹⁶. Nevertheless, Letzel

et al advocate to take a rather indulgent stand against the Ginkgo efficacy trials conducted in the past. They give the following arguments¹⁶:

1. Dementia and its related disorders are differently assessed - both conceptually and in health care practice - in (continental) Europe, where the significance of mild forms of dementia for those affected is often underestimated, and where the vast majority of *Ginkgo biloba* intervention trials have been conducted, and in the USA, where Alzheimer's disease is the main focus of research. While most of the studies on *Ginkgo biloba* extract were influenced by European requirements, the planning of tacrine studies, for instance, was subjected to the American 'school of thought' on dementia and the corresponding standards.
2. Since the nootropic effect of *Ginkgo biloba* is claimed not to be specifically associated with one particular pathogenetic route (several mechanisms of action are assumed, which would justify a 'broad spectrum' application), previous studies on this active substance included most of the times a mixture of AD-patients and patients with vascular dementia. From a therapeutical point of view, the need to separate AD- and VD-patients is not urgently felt by health care practitioners in many European countries. Trials with tacrine, on the contrary, have been carried out exclusively in patients with AD, since tacrine - a cholinesterase inhibitor - is supposed to counteract specifically against the cholinergic deficit which is present in AD. As a consequence, Ginkgo and tacrine trial populations consist of patients with cerebral function impairment caused by different factors, depending on the active substance studied.
3. On the other hand, this observed difference in the composition of the study populations between various anti-

dementia drug trials is to a certain extent artificial. In most of the tacrine studies the diagnosis AD was based on the criteria developed by the National Institute of Neurological and Communicative Disorders and Stroke - Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA criteria)³³. However, patients with AD, VD, et cetera, cannot be distinguished by clinical means alone. Given the scientific knowledge available at present, a certain diagnosis of AD can only be obtained post mortem. Therefore, the inclusion diagnosis according to NINCDS-ADRDA criteria is always 'probably Alzheimer's disease', and some of these patients may turn out to have no AD, but VD or something else. If one further recognizes that AD is substantially more common than VD (about 50% versus 20% of all cases), the patient populations with mixed diagnoses and with probable AD are not completely different. In addition, when we study the efficacy of nootropics, aetiopathogenic differentiation between AD and VD can currently be ignored without a substantial loss of informative value. Differentiation would only be necessary if the specific pathogenic efficacy profile of a substance were to be investigated.

4. Although it was not explicitly diagnosed, former *Ginkgo biloba* trials must have contained a substantial number of cases with dementia, be it AD or VD. Wasting scarce information would be the consequence of neglecting this. Moreover, a distinction should be made between trials based on hospitalized patients and trials based on out-patients, as sophisticated diagnostic tools are usually more close at hand in an institutional setting.

The diagnostic uncertainties touched upon above, may seriously hamper the realization of both a criteria-based systematic re-

view and a meta-analysis based on statistical pooling.

3.1.3 Outcome measurement uncertainties

Analogous problems to those complicating the diagnosis of dementia and related disorders manifest themselves in the outcome measurement domain. To evaluate the efficacy of Ginkgo treatment the trials conducted in the past have relied on a wealth of different outcome measures, covering various levels of functioning (general well-being, cognitive functioning, daily life activities, psychopathology, psychometric capacities, clinical signs and symptoms, electrophysiological reactions, et cetera). Most of the trials have incorporated a variety of outcome measures without a predefined hierarchy. Again, the former researchers can hardly be criticized for this. Only quite recently stricter requirements have been formulated for the evaluation of nootropics and anti-dementia drugs, as far as the outcome measurement and evaluation of efficacy of these drugs are concerned⁴¹⁻⁴⁵. These regulations and guidelines are being revised and adapted periodically. They state, for instance, that the outcome measures and the success criteria derived from these outcome measures should be chosen in advance of the trial onset, i.e. should be already included in the trial protocol. Success criteria should not only be formulated in terms of statistical significance (hypothesis testing), but also in terms of clinical relevance (meaningfulness of the absolute changes and differences in parameter values). The assessment of the ability of interventions to slow down, stabilize, or improve cognitive impairment should be investigated at least at three complementary levels: 1. cognitive functioning (psychometric testing); 2. psychopathology (presence of clinical signs and symptoms); 3. behaviour (influence of cognitive functioning on daily life activities, social functioning, etc.).

Moreover, it is required to make a distinction between primary (= main) and secondary outcome parameters, and, connected with this, between a confirmatory and a descriptive analysis of the data collected during a trial. A responder analysis should be considered as well. A recent review, comparing the designs and results of *Ginkgo biloba* special extract trials with those of two other active substance classes (tacrine, nimodipine) suggested that only a few recent trials have succeeded in keeping pace with the most current standards¹⁶. Again, this circumstance seriously hampers the conduct of a systematic review.

3.2 STRATEGY FOR SEARCHING THE LITERATURE

The following sources of information were consulted to identify relevant Ginkgo trials:

1. Searching of electronic bibliographic databases: Medline, 1966-1997 (National Library of Medicine; Index Medicus; by means of SilverPlatter 3.0; emphasis on publications in medical journals); Embase, 1974-1997 (Excerpta Medica; emphasis on publications in medical journals); PsycLIT, 1972-1997 (by means of SilverPlatter 3.0; emphasis on publications in the behavioural sciences); Current Contents, 1997-1998 (on line, by means of CC Search (R)).
2. Systematic hand-searching of conference proceedings and abstracts relevant to dementia (dementia, geriatric, psychiatric, psychogeriatric, neurology, and stroke conferences) or phytotherapy.
3. Checking the contents of several journals (hand-searching), especially Germany and France based medical journals.
4. Checking references extensively in review articles, in original articles on clinical research, and in textbooks.
5. Personal communications with acknowledged Ginkgo investigators and with the

major pharmaceutical companies involved in manufacturing and trading *Ginkgo biloba* preparations: Dr. Willmar Schwabe Arzneimittel GmbH (Karlsruhe, Germany), Lichtwer Pharma (Berlin, Germany), and IPSEN (Paris, France). These companies seem to be involved in most of the research activities in the Ginkgo field.

The free-text search of the electronic bibliographies Medline, Embase and PsycLIT was based on a series of keywords which referred to the therapeutic intervention of interest: Ginkgo, gingko, ginkco, ginko, gingho, tebonin, tebonine, kaveri, tanakan, rōkan, egb, egb-761, ginkobene, Ginkgolip, Ginkgold, Ginkgolic, Ginkgoextrakt, gincosan, ginkoba, Ginkgolide, Ginkgolide-b, Ginkgoliden, Ginkgolides, Ginkgolids, Gingkolide, Gingkolides, Ginkolide, bilobalid, bilobalide, bilobalids.

3.3 STRUCTURE OF THE TABLES IN THE CURRENT REVIEW

The structure of the tables that we made to summarize the design and the results of the intervention trials retrieved for the current review, will be explained in this section. We have organized the information in each of the six extended tables under 8 different headings.

Column 1

Column 1 shows the names of the author(s) and the year of publication of the trial report. Within each table the trials are arranged in a chronological sequence (year of publication).

Column 2

Column 2 shows the key information regarding the patient population in each trial: the diagnostic criteria and a summary of the main disease characteristics. Some additional in-

formation on the composition of the study population may be listed in column 8 ("Remarks").

Column 3

Column 3 shows the total number of randomized study subjects (first row), as well as the numbers allocated to each of the intervention subgroups (second row). For cross-over trials the second row remains empty, as each study participant is exposed to both the experimental drug and the reference drug. If the number of subjects included in the analysis differs from the number randomized, e.g., due to dropping out, the total and treatment-specific numbers of analysed persons are displayed on the third and fourth rows of the same field, provided that this information could be retrieved from the trial report. The analysis may have been based on either the intention-to-treat (ItT) or the per-protocol (= valid-cases) principle. For some studies we have entered additional information into column 8 of the table. Some trials had a multi-comparison design, e.g., they contained both a placebo control group and a reference drug control group, in addition to the Ginkgo group. In such cases only the information concerning the verum-placebo part of the trial will be presented. In this column as well as in other columns the notation 'n.d.' (= not described) is used to indicate that the relevant information could not be retrieved from the published data.

Column 4

Column 4 provides information regarding the age and gender composition of the study population. The age range of the persons admitted to the trial and/or the mean (or median) age for the undivided study population are specified in the first row, the gender ratio, i.e. the numbers of male and female participants, respectively - preferably in the randomized population, and otherwise in the analysed population - are shown in the sec-

ond row of this column. Occasionally, information regarding the distribution according to some other very important group characteristic, e.g., Alzheimer's dementia vs vascular dementia, is entered under this heading as well.

Column 5

Column 5 deals with the study treatment: product name (brand name and/or special extract code name), route of administration, daily dosage and duration of the intervention. *Ginkgo biloba* extract can be administered in various ways. Most of the times it is supplied per os, either in a solid form (film-coated tablets, dragees, capsules), or as a solution (drops). Intravenous or intramuscular injection, and intravenous infusion are alternative options, especially in clinical settings. Tablets and solutions may contain various amounts of the extract that encompasses the active principles. Eventually, the effect is assumed to be dependent on the total daily dose and its time-related distribution (number and timing of intakes). Regular dosage schedules for Tebonin[®] tablets - Tebonin[®] is one of the most popular Ginkgo brand names - are, for instance, 2 or 3 film-coated tablets a day (1 tablet = 40, or 80, or 120 mg of active substance (dry extract; input : output ratio = 50 : 1)). In the past a total daily amount of 80-160 mg special extract, usually corresponding with 1-4 tablets, was regarded to be sufficient, but recently a tendency to elevate this dose a little bit can be noticed, especially within the context of clinical research.

To indicate the level of exposure to the study medication, we have aimed at reporting the quantity supplied on a daily base, expressed in mg of active substance. If this figure was not available, we had to rely on the next best information (e.g., ml of solution per day or number of drops per day). The dosage scheme is not further specified (e.g., time-related distribution of the daily dose).

The last row shows the duration of the intervention, covering the period from onset (baseline-measurement, after randomization) until the time of the final outcome measurement. Run-in phase and wash-out period are not included in this estimation of the length of exposure. Prolonged follow-up after completion of the randomized part of the trial (e.g., an 'open' trial) has not been taken into consideration either. In case of a crossover trial the period of Ginkgo treatment is presented. The length of the treatment period is expressed in either hours (h.; so, registration of an acute intervention effect), days (d.), weeks (w.), or months (m.). The intervention contrast - GBSE vs either placebo, or a reference drug, or no treatment at all - is not specified in this column, but in column 8. In a few studies Ginkgo treatment acted as adjuvans therapy to another therapy. If this was the case, the standard therapy is mentioned as well.

Column 6

Column 6 contains information on the outcome measures in each study. A wealth of different endpoints have been chosen to detect Ginkgo effects. These outcome measures cover several levels and domains of achievement: global assessment of (change in) health and cognitive state; assessment of clinical signs and symptoms (presence, severity); assessment of functional signs and symptoms; cognitive rating scales; measurement of functional capacity under laboratory conditions (psychometric and physical testing, e.g., aspects of memory function, attention, reaction time, specific parameters of information processing in the brain, concentration, coping with interference, visuo-spatial orientation); neurophysiological and -anatomical testing (e.g., EEG-patterns, electronystagmography, craniocorpography, CCT); measurement of functional capacity in daily life; and behaviour.

Only a minority of the study reports communicated an explicit hierarchy of endpoints. If so, this structure is conserved in the tables, for instance by distinguishing primary and secondary outcome measures (printed in italics). Otherwise, only the most relevant data are selected for presentation, according to the following priority setting: global assessment; general well-being, quality-of-life assessment; behaviour, daily life activities and functioning; psychometric and neuropsychological test results (functional capacity under laboratory conditions); clinical signs and symptoms; neurophysiological test results; physical signs and symptoms; and physiological and biochemical laboratory values.

With regard to the questionnaires, clinical rating scales and psychometric tests, the tables provide information on both the generic type and the specific format ('brand name') of the measurement tool concerned (e.g., depression, Hamilton Depression Scale). In order to facilitate the interpretation of the outcomes recorded, we have added for the most frequently used outcome measures some information on the scale or test characteristics, such as dimensions and subscales, and the range of scores (underlining of one extreme value indicates the most favourable result). For subjective outcome measures the source of information is specified (e.g., patient, physician, caregiver, relatives, investigator).

In addition to the efficacy much information has been collected regarding the safety and tolerance of the pharmacological agents evaluated in most of the trials, especially the more recent ones. This was often done in a systematic and extensive fashion (monitoring of adverse events, serious adverse events). The same holds for information on compliance and adherence to the treatment protocol, withdrawal, missing values of outcome parameters, etc. We have chosen to withhold his kind of information from the ta-

bles. In general, one agrees that GBSE is safe and tolerable. Almost all studies failed to establish a relationship between study treatment and adverse events. If such a relationship could not be fully excluded, rather minor health consequences were the main concern.

Column 7

Column 7 summarizes the reported results for the outcome measures mentioned in the corresponding rows of column 6. Unless stated otherwise, only information related to the final outcome measurement time-point is presented. 'Vs' is inserted to contrast the results registered for the Ginkgo group (mentioned first) and the reference group (either placebo group, or alternative drug group). In the case of a continuous outcome measure the numerical results may refer to the mean (or median) status scores at the end of the intervention period. They may also refer to the mean (or median) changes in status score from baseline to end. When the change in status over time is the outcome parameter of primary concern, the numerical value is preceded by either '+' or '-' to indicate improvement and deterioration, respectively, for the relevant comparison group. In the case of a discrete (categorical) outcome measure, most of the time '%' has been used to indicate the relative size of each category. Sometimes, especially for small-sized trials, absolute numbers have been entered to avoid sham-precision.

For some trials numerical results could only be generated via extrapolation of the information stored in a graphical format. This may have resulted in slightly imprecise outcome estimates.

If retrievable from the trial report, the results of formal hypothesis testing (P-values) have been added to the numerical outcome values, in order to show the statistical significance of each effect. We have just copied this information, without further information

on the test statistic used and without checking the appropriateness of the statistical data processing.

Column 8

The final column provides some additional, miscellaneous information on the trial design and the trial results. First of all this column contains information on the study design and the main intervention contrast. The third row sheds some light on the source of recruitment of the patient population: out-patient versus hospitalized, type of medical practice, etc. The rest of the information is optional and not systematically entered for each trial (dependent on relevance, etc.).

4 OVERVIEW OF TRIALS ON THE EFFICACY OF *GINKGO BILOBA* IN CEREBRAL INSUFFICIENCY

4.1 CATEGORIZATION OF TRIALS

Since 1975 a considerable number of trials have been published which studied the efficacy of *Ginkgo biloba* special extract in subjects with various manifestations of cerebral insufficiency. We have entered 55 in the current review. Although all dealt with the same, rather vague concept of 'cerebral insufficiency', they tend to show more differences than similarities. The recent review of Letzel et al regarding the nootropic effect of *Ginkgo biloba* in comparison with two other active substance classes (tacrine, nimodipine) illustrates various approaches to distinguish a limited number of distinct, more homogeneous subcategories of studies¹⁶. It shows that reference can be made to various design aspects, such as the diagnostic inclusion criteria applied, the measures used to assess the severity of the disease, or the type of instrument used to assess the intervention outcome.

To structure the current overview we have divided the studies into the following six categories, based on the main treatment indication:

1. Studies focussing on patients with dementia, including Alzheimer's disease, vascular dementia, and mixed type of dementia
2. Studies focussing on non-demented patients with cognitive decline
3. Studies focussing on patients with cerebral insufficiency expressed primarily by physical, non-cognitive manifestations (tinnitus, vertigo, hearing loss, etc.)
4. Studies focussing on patients with a mix of cognitive and non-cognitive symptoms, diagnosed and evaluated with a standard symptom checklist.
5. Studies focussing on patients with affective manifestations of cerebral insufficiency, especially studies focussing on depressive mood as the leading symptom
6. Studies focussing on patients with cerebral insufficiency according to diagnostic information that does not permit further differentiation of the disease state (cerebral insufficiency, unspecified).

We are aware that our choice of structuring the studies can be challenged and that some trials do not fit exactly in one of the categories.

In discussing the empirical evidence we pay special attention to the first category of studies. In our opinion these trials are of special interest, given their focussed character - restriction of the patient definition to a relatively precisely described and homogeneous disease - in combination with the claim that *Ginkgo biloba* special extract might benefit both patients with Alzheimer's dementia and patients with vascular dementia. Moreover, the majority of these studies in this group seem to surpass most of the trials in the other subgroups as for methodological quality.

4.2 SUMMARY OF STUDY CHARACTERISTICS AND STUDY RESULTS

4.2.1 Dementia

We found 9 RCTs which evaluated the effect of Ginkgo treatment in patients diagnosed with dementia. However, only 5 of these have been designed and conducted in agreement with the gradually evolved and currently accepted scientific guidelines and prescriptions for the diagnosis of dementia and the evaluation of therapy-induced changes in the state of dementia⁴⁶⁻⁵⁰. According to these guidelines and prescriptions, the patient selection should be based on the DSM-III-R / DSM-IV, the ICD-10, or the NINCDS-ADRDA³³ criteria for dementia. And the outcome assessment should be based on three different levels of achievement: objective psychometric performance, clinical psychopathology (both global and specific), and (cognition-related) daily life activities. It is not surprising that only trials completed during the 1990s fulfil these criteria. We decided to deal here also with trials which claim that they have included patients with dementia, although the appropriateness of this claim cannot be verified from the documentation available. Moreover, one should recognize that many of the trials allocated to some of the other headings (e.g., cognitive impairment) probably have harboured a considerable number of demented patients as well, who have been wrongly categorized due to the deployment of insufficient diagnostic equipment.

Details of the dementia trials are communicated through table 1 (see end of this chapter).

Hofferberth performed a RCT in 40 hospitalized patients with incipient dementia of the Alzheimer type⁴⁶. All patients (age range: 50-75 years) were treated during 3 months

with either Tebonin® forte (tablets, 80 mg/day) or placebo. The Syndrom Kurz Test (SKT) of memory and attention (scale: 0-27), was the primary outcome measure in this trial. Ginkgo treatment turned out to be beneficial, according to the observed mean change in SKT-score over 3 months (+5 vs -2 points; $p < 0.001$). Fifty percent of the Ginkgo users and none of the placebo users showed an improvement exceeding 5 SKT-points. A global assessment by the physician, the Sandoz Clinical Assessment-Geriatric Scale (SCAG; measuring the degree of clinical psychopathology), and various EEG characteristics, were some of the secondary outcome parameters in this study. The patients in the verum group showed more favourable scores also on these parameters compared to the patients in the placebo group.

Kanowski et al conducted a placebo-controlled, parallel trial in 216 out-patients with Alzheimer's dementia or vascular dementia (both in a mild to moderate state), who were recruited from 39 centres (practices of general practitioners, neurologists, psychiatrists, internists)⁴⁷. After a 4-week run-in phase, the participants were randomly allocated to a 24-week treatment with either EGb 761 (240 mg/day) or placebo capsules. 205 Patients were involved in the intention-to-treat analysis. However, the investigators preferred an analysis based on the 156 participants who completed the trial (79 Ginkgo, 77 placebo; 51 male, 105 female; 31 AD, 125 VD) as the confirmatory analysis of the trial results. The SKT (range of scores 0-27), the clinical global impression of change by the physician (CGI-2, range of scores 1-7), and the Nuremberg Geriatric Observation Scale (NAI-NAB) - a rating scale which measures behaviour indicating the level of coping with everyday tasks and the level of independency of care (rated by relatives of the patient) - were applied as the primary outcome

measures. The mean change in the SKT total score was +2.2 for the Ginkgo group vs +0.8 for the placebo group. With a decrease of at least 4 points on the SKT-scale acting as the criterion for success, 38% of the Ginkgo users vs 18% of the placebo users turned out to be responders ($p < 0.05$). Thirty-two percent of the Ginkgo users and 17% of the placebo users were found to respond according to the CGI-2, when the categories 'much improved' and 'very much improved' were used as the success criterion. The mean changes on the NAI-NAB were +0.9 vs +0.6 points for the Ginkgo group and the placebo group, respectively, with 33% vs 23% responders ($p < 0.1$; response criterion ≥ 2 NAI-NAB points). An overall responder analysis - criterion: response for at least 2 out of the 3 primary outcome measures - yielded 28% vs 10% ($p = 0.005$) responders among the Ginkgo users and the placebo users, respectively.

Haase et al reported on the results of a parallel, placebo-controlled trial in 40 out-patients of a neuro-psychiatric practice⁴⁸. The study subjects received an intravenous treatment, with 200 mg EGb per day for the Ginkgo group, sustained over a 4-week period. The daily dose was administered by infusion, 4 days a week. Included in the trial were patients with a moderate state of dementia, either AD, or VD, or a mixed form. The Kurzttest für Allgemeine Intelligenz (KAI), the clinical global impression of change by the physician (CGI-2), and the Nuremberg Geriatric Observation Scale (NAI-NAB) were the primary outcome measures in this trial. With respect to the NAI-NAB, 17 Ginkgo users vs 3 placebo users improved (mean change in score: +3.6 vs +0.3 points ($p < 0.01$)). With regard to the CGI-2, 17 vs 7 participants improved ($p < 0.0001$). The KAI-test for information processing capacity resulted in an improvement for 7 vs 1 of the participants ($p < 0.03$).

Maurer et al studied the effect of treatment with Tebonin® forte (tablets, 240 mg/day, intervention sustained during 3 months) in patients suffering from a not too advanced stage of Alzheimer's dementia^{49,51}. Eighteen subjects were included in the analysis. The Syndrom Kurz Test was chosen as the outcome measure of primary interest. The changes in SKT total score were +2.9 vs -0.8 ($p = 0.013$), departing from baseline scores amounting to 19.7 vs 18.1 (scale range 0-27). Positive, though not statistically significant effects were also reported for the other outcome parameters: the number-connection test (NAI-ZVT-G), the ADAS cognitive and non-cognitive subscales, the clinical global impression of change (CGI), and several electrophysiological parameters (AEP-300, EEG-topography). 8 Out of 9 Ginkgo users and 5 out of 9 placebo users stated that they would like to continue the treatment⁴⁹.

The first clinical trial on the efficacy of *Ginkgo biloba* special extract in the USA was reported by Le Bars et al⁵⁰. Included were 327 persons (aged 45-90 years) from 6 out-patient centres, with a mild to moderately severe state of dementia according to the DSM-III-R and ICD-10 criteria, either AD or VD. After a 2-week run-in period the patients were randomly allocated to either Ginkgo (tablets, daily dose of 120 mg) or placebo treatment, for a period of one year. 78 Patients in the Ginkgo group and 59 in the placebo group completed the trial. 309 Subjects could be entered in the intention-to-treat analysis, whereas 202 patients met the criteria for the per-protocol analysis. Primary outcome measures in this trial were: 1. the cognitive subscale of the Alzheimer Disease Assessment Scale at the cognitive function level (ADAS-Cog; performance-based (memory, language, praxis, orientation); 11 items, yielding a 0-70 sum score range); 2. the Geriatric Evaluation by Relative's Rating

Instrument at the daily living and social behaviour level (GERRI; rated by the caregiver after 14 days of observation; 49 items, divided over 3 subdomains: cognition, social functioning, and mood; item-, subscale- and total scale score range: 1-5); and 3. the Clinical Global Impression of Change (CGIC; scored by the clinician; scale: 1-7). The mean change in ADAS-Cog score showed less impairment for patients allocated to the Ginkgo group than for patients in the placebo group (-0.1 vs -1.5 ($p=0.04$)); 50% Vs 29% of the patients in each of both groups improved with ≥ 2 points. The mean change of the GERRI-overall score was also more favourable for the Ginkgo treatment group: +0.06 vs -0.08 ($p=0.004$). Improvement was registered for 37% and 23% of the Ginkgo and placebo groups, respectively. The treatment groups did not differ with respect to the CGI-scores (4.2 vs 4.2 ($p=0.77$)). The analysis did not reveal any effect modification by AD/VD-subgroup.

The four remaining trials in this category, which used more relaxed criteria to identify patients with dementia, were all designed as parallel, placebo-controlled trials^{31,52-56}. The experimental treatment consisted of 120-150 mg of active substance per day, for approximately 3 months.

Weitbrecht & Jansen reported a beneficial effect of Ginkgo treatment in patients with a mild to moderate form of primary degenerative dementia, according to both various clinical rating scales (hetero-assessment of general health state, Crichton Geriatric Scale, SCAG) and various psychometric tests (figure-symbol test, number-repeat test, flicker fusion frequency, reaction time)^{52,53}.

Halama evaluated the efficacy of Ginkgo treatment in 50 subjects with cerebral insufficiency classified as either degenerative dementia (Alzheimer's disease) or vascular dementia^{54,55}. The Ginkgo cohort showed

better results for the Syndrom Kurz Test (SKT; +4.2 vs +1.3 points), the global assessment of treatment efficacy, and the number-connection test. 7 Out of 11 rated clinical symptoms of cerebral insufficiency showed a significantly larger number of improvements among Ginkgo users than among placebo users.

Hartmann & Frick studied the effect of *Ginkgo biloba* in 52 patients with vascular dementia. Patients presenting at least one symptom out of four symptom categories that are typical for the so-called organic brain psychosyndrome were considered eligible for the trial. A statistically significant effect was only recorded for the Grünberg Verbal Memory test, not for the other outcome parameters (global assessment of efficacy, 12 clinical symptoms, trail making test)³¹.

In a trial with 90 subjects, Vesper & Hänsgen reported a positive effect in favour of the Ginkgo group on the CI-scale of subjective cerebral insufficiency troubles, both during the first part and the second part of the intervention trial. The rating of changes in each of 8 mental and physical functions on a 1-7 point transition scale (4 indicating 'no change', 7 'very much improvement') pointed to a beneficial effect as well; the scores on the 8 functions varied from about 4.7 to 5.0 in the Ginkgo group and from about 3.9 to 4.2 in the placebo group. A series of computer-aided psychometric tests also suggested a positive effect of Ginkgo treatment. No difference was found for the Short Test of General Intelligence (KAI)⁵⁶.

We conclude that the published randomized clinical trials which have been designed to prove the efficacy of *Ginkgo biloba* special extract in demented patients, show remarkably consistent results. All trials show positive, often statistically significant effects in patients subjected to Ginkgo treatment compared with patients under placebo treatment, for most of the outcome parameters as-

sessed. One may argue, however, that the reported effect sizes are less impressive and less convincing from a clinical point of view. How should one value, e.g., a mean gain of 1.4 points for Ginkgo over placebo treatment on the SKT-scale of memory and attention (maximum improvement: 40 points)⁴⁷ or on the ADAS-Cog scale (maximum improvement: 70 points)⁵⁰?

4.2.2 Cognitive impairment

The second category consists of 8 RCTs that focussed on patients with cerebral insufficiency manifested by cognitive impairment probably not being dementia^{30,57-63}. Four studies were conducted in France, two in the UK, one in Germany, and one in Austria. Again, it should be emphasized that the reported diagnostic criteria do not always preclude the presence of dementia⁶³. Several terms and criteria have been proposed to indicate a type of cognitive decline that can be distinguished from dementia as a separate pathological entity. Age-associated memory impairment (AAMI)^{39,64}, age-associated cognitive decline (AACD) and cognitive decline no dementia (CIND) are some of the most familiar names. According to the provided patient recruitment information, only one trial strictly adhered to the AAMI-criteria⁶².

Table 2 presents the main features of these cognitive impairment trials (see end of this chapter).

In addition to 7 parallel trials - 6 of them placebo-controlled, and 1 with nicergolin (NCG) as the reference substance⁶⁰ - we identified 1 cross-over trial⁶¹. One of the parallel trials had a factorial design, with memory training as the second intervention of interest⁵⁸. The most frequently occurring daily dose was 120 mg/day, and treatment duration varied from 2 to 12 months. The cross-over trial aimed at demonstrating the acute effect of a large single dose of EGb 761.

The outcome measures used in these trials included the global assessment of the intervention effect (4 studies); various clinical rating scales (e.g., MMSE in 2 studies, Global Deterioration Scale); quality-of-life assessment; depression; self-assessed well-being and behaviour; hetero-assessed functional capacity; memory and attention; level of drug prescription; and finally psychophysiological parameters (e.g., EEG brain-mapping). Except for two studies^{60,63}, a clear hierarchy of outcome measures was not given.

Chartres et al found no difference in global effect between Ginkgo and placebo, a small gain in the level of cognitive functioning (MMSE, Reisberg's GDS, Plutchik GRS) for the Ginkgo users, no difference in the development of several clinical signs, but a remarkably larger decline in the prescription of (other) psychotropic drugs for the Ginkgo group⁵⁷.

Israel et al observed an extra effect of Ginkgo, on top of the effect of memory training. Within the subgroup that did not receive memory training, the Ginkgo users felt more satisfied than the placebo users. Part of the results of the memory test battery (long-term memory, fluency test) was in favour of Ginkgo as well⁵⁸.

Wesnes et al failed to detect a significant overall effect of Ginkgo. They also found no difference in actual daily activities between Ginkgo and placebo users. However, the Ginkgo group performed better on most of the psychometric tests, which were in this trial at the forefront as outcome measures³⁰. In the intervention trial of Rai et al (computerized) psychometric testing was also strongly emphasized. Compared with the placebo group, the Ginkgo group performed better on most of the tests. A positive effect, be it small, was also found for the MMSE⁵⁹.

Franco et al reported better results for Ginkgo than for nicergolin on almost all outcome measures⁶⁰.

The cross-over trial conducted by Allain et al evaluated the acute effect of two different doses of Ginkgo and placebo through a Latin square design. Series of drawings and words were shown, with presentation times varying. The number of correctly recalled objects was registered. It is a well-known fact that drawings can be remembered more easily than words. When in this trial for each of the presentation times the difference between the numbers of correctly recalled drawings and words was calculated, it turned out that the breakpoint and the 'dual coding' point that marks the onset of a significant difference between word recall and drawing recall, had been shifted to a shorter presentation time under influence of Ginkgo treatment, irrespective of the treatment dose. This suggests an increase of the relative speed of information processing due to Ginkgo treatment⁶¹.

Semlitsch et al assessed the influence of Ginkgo in AAMI-patients by means of psychometric tests, EEG brain-mapping, a checklist expressing subjective thymopsychical well-being, and neurophysiological examinations reflecting the quality of cognitive information processing. Acute (3 hours after the first dose, week 0), chronic (before the last dose, week 8 of treatment), and superimposed (3 hours after the last dose, week 8 of treatment) effects of Ginkobene® on P300 latency were registered. The shortened latency may reflect a decrease of the stimulus-evaluation time, evoked by Ginkgo treatment⁶².

Oswald et al evaluated the effect of *Ginkgo biloba* within three domains of performance: fluid cognitive performance (objective psychometric testing), subjective well-being (self-rating), and functional performance (hetero-assessment). Only at the fluid cognition level (labyrinth test) a statistically significant effect of Ginkgo was found. A responder analysis which covered all three domains of

performance did not yield a significant result⁶³.

We conclude that the results of the trials discussed under this heading point to a beneficial effect of Ginkgo treatment, although not unequivocally and not for all outcome measures.

4.2.3 Non-cognitive signs and symptoms

We identified 12 randomized clinical trials which had their main focus on the treatment of 'physical', non-cognitive signs and symptoms of cerebral insufficiency, such as vertigo (dizziness), tinnitus (ear noise), hearing loss, cochlear deafness, and ataxia (table 3; see end of this chapter).

These studies were reported from France (6), Germany (5), and Sweden (1)^{25,27,65-75}. A cross-over design, contrasting Ginkgo treatment with cinnarizin (CIN) and placebo, respectively, was employed in two trials^{66,73}. The other studies were designed as parallel trials. Six of them contrasted *Ginkgo biloba* with placebo, the remaining four compared Ginkgo with another active compound (nicergolin (NCG), almitrine-raubasine (ARB), naftidrofuryl). Four studies focussed on vertigo as the leading symptom, also 4 on tinnitus, 2 on hearing loss, and 2 had a mixture of patients with vertigo, tinnitus and hearing loss. The number of patients varied from 20 to 80 in most of the trials, and exceeded 100 in only two trials^{25,72}. The mean age of the study population was close to 50 in the majority of the studies. The 'average' intervention scheme consisted of a daily Ginkgo dosage of 120-160 mg/day, administered as an oral solution during 1-3 months. In one study Ginkgo extract was applied as a supplement to a physical vestibular training programme⁷⁰. In 2 studies only subjects were admitted who had experienced a positive effect of Ginkgo treatment during a preceding, open trial^{67,73}.

The following endpoints were used in these trials: global assessment of the treatment effect^{25,27,65,66,71-73,75}; assessment of functional signs and symptoms through a questionnaire, VAS-scale, or clinical rating scale^{25,27,65,67-70,72-75}; and various forms of paraclinical otoneurological examination, e.g., vocal and tonal audiometry, electronystagmography, craniocorpography (CCG) and Romberg's test (posturographic body sway).

The trials which contrasted Ginkgo treatment with a reference drug, may be considered less informative with regard to the potential efficacy of Ginkgo. Nevertheless, it is remarkable that in most of these trials Ginkgo surpassed the reference drug for nearly all the parameters evaluated^{65,66,71,72}.

In the trial reported by Schwerdtfeger the overall conclusion, based on a questionnaire, electronystagmographic evaluation, and several other neuro(oto)logical tests, pointed to a beneficial effect of Ginkgo⁶⁷. Claussen found a positive effect on both parameters evaluated: subjective dizziness symptoms and lateral body sway^{68,69}. Hamann reported a positive effect on the degree of posturographic body sway, whereas the change in subjective dizziness complaints was almost equal for the Ginkgo group and the placebo group⁷⁰. Meyer reported beneficial effects of Ginkgo, which were (borderline) statistically significant for all the relevant endpoints (global assessment of change in clinical signs, change of intensity of tinnitus complaints, change of level of discomfort due to tinnitus, time needed to reach significant improvement or disappearance of complaints)⁷². Statistically significant results for all outcomes measured were also communicated by Haguenaer et al²⁷.

Morgenstern & Biermann found a larger regression of tinnitus loudness, assessed by means of tonal audiometry, for the Ginkgo

group than for the placebo group. However, this result was not accompanied by a positive effect of Ginkgo on the patient's subjective impression of the change in tinnitus complaints⁷⁵.

In contrast with these studies the placebo-controlled cross-over trial published by Hölgers et al did not reveal any positive effect of Ginkgo in patients suffering from tinnitus⁷³.

We conclude that almost all trials included in this review suggest that Ginkgo is beneficial for patients with vertigo, tinnitus and comparable manifestations of cerebrovascular insufficiency, with only one trial out of tune.

4.2.4 Subjective cognitive and non-cognitive symptoms

Table 4 (see end of this chapter) shows the details of the subcategory of 5 randomized clinical trials - all reported from Germany - which focussed on the treatment of patients with cerebral insufficiency, manifested by a combination of both cognitive and non-cognitive clinical symptoms^{23,24,29,76,77}.

A common feature of these trials is that the investigators applied a symptom checklist to assess the presence and severity of 11 or 12 clinical symptoms, both for the patient selection and for the outcome evaluation. Memory loss, forgetfulness, lack of concentration, dizziness, tinnitus, fatigue, headache, loss of performance, lack of motivation, depression, anxiety, and sometimes confusion, made up a combination of symptoms, which were scored on an ordinal response scale that expressed the degree of nuisance caused by each symptom (absent, mild, moderate, severe). In one study a slightly different symptom checklist was used⁷⁶, and in another study the primary emphasis was on outcome measurement through psychometric and neurophysiological testing (Vienna Determination Test, number connection test, saccadic eye movements, EEG-analysis,

evoked potentials)⁷⁷. The same symptom checklist also played a prominent role in two other trials, which we discussed previously under the heading 'anti-dementia trials'^{31,54,55}.

All studies in this category were placebo-controlled trials. The size of the study population varied from 50 to more than 300 participants, the daily dose of *Ginkgo biloba* special extract from 120 to 160 mg/day, and the duration of the intervention from 4 to 12 weeks. As table 4 demonstrates, most of the symptoms were found to react (significantly) better to Ginkgo than to placebo treatment. Global assessment of the treatment effect in addition to the administration of the symptom checklist, either by the physician or by the patient, always proved to be in favour of Ginkgo. Hofferberth reported a positive effect of Ginkgo for the psychometric and the neurophysiological parameters as well⁷⁷.

We conclude that Ginkgo treatment proved to be effective in all trials in this category.

4.2.5 Depression as the leading symptom of cerebral insufficiency

We found 3 randomized clinical trials that concentrated on depressive mood as the major cerebral insufficiency symptom of interest⁷⁸⁻⁸⁰. Details of these trials are presented in table 5 (see end of this chapter).

In each of these trials Tebonin[®] forte (120-240 mg/day during 8-12 weeks) was compared with placebo treatment. In two studies the patients had turned out to be resistant to previous therapy with (other) antidepressants^{78,80}. In both trials EGb 761 was administered in combination with other antidepressants^{78,80}. Zung's Self-Rating Depression Scale⁷⁸ and the Hamilton Depression Scale^{79,80} were used as the primary outcome measures.

A beneficial effect of Ginkgo treatment was reported in all 3 trials, viz. 5.3 points improvement vs 5.2 points deterioration on Zung's scale⁷⁸, 6.0 vs 4.6 points improvement on the Hamilton Depression Scale (HAMD)⁷⁹, and 9.5 vs 1.0 points improvement, also on the HAMD-scale ($p < 0.01$)⁸⁰, respectively. In the study of Schubert & Halama, 75% of the Ginkgo users vs 30% of the placebo users showed a ≥ 5 point decrease on the HAMD-scale⁸⁰. In the study of Halama a positive effect of Ginkgo treatment was reported for an additional primary outcome measure (late acoustic-evoked potential) and for a series of secondary outcome measures (Cerebral Insufficiency scale, Cerebral Insufficiency subtests, Kurztest für Allgemeine Intelligenz (KAI), subjective assessment of depression and concentration)⁷⁸. No significant differences were registered for the secondary outcome measures in the other two 'depression' trials, except for a better result on the KAI for the subgroup under Ginkgo treatment and a larger decrease in the number of patients with moderate or severe symptoms of depression and concentration based on subjective assessment by the physician, in the study of Schubert & Halama⁸⁰.

We conclude that the trials in this category consistently suggest a favourable effect of Ginkgo on depressive mood.

4.2.6 Cerebral insufficiency, unspecified

We reserved a rest category for 18 randomized clinical trials in the field of cerebral insufficiency which were published without sufficient diagnostic information to enable categorization to a more specific subgroup. Most of these trials date from the 1970s and the first half of the 1980s. Nine were conducted in Germany, 6 in France, and 3 in Italy^{26,28,81-97}.

As table 6 shows, 14 trials have been designed as placebo-controlled parallel trials, 2 as parallel trials contrasting Ginkgo treatment with a (vasoactive) reference drug (vincamine, dihydroergotoxin)^{87,95}, and 2 as placebo-controlled cross-over trials^{85,91,92}. The number of participants amounted to 30 and 90 in the cross-over trials, and varied from 14 to 189 in the parallel trials. The average age of the participants approached 70 years. In the majority of the trials a daily amount of 120 mg of active substance was administered. A higher exposure level was chosen in 5 trials (150-160 in 4 trials, 240 in 1 trial). The most common mode of administration of the study medication was per os - in 8 trials as a solution, in 5 trials in tablet form - while in 1 study the active substance was administered intravenously for 15 days⁸⁷; the reports of 4 other studies lack a clear description of the mode of administration. In most trials the duration of the intervention was 1-3 months, in two studies 6 months, and in one study 12 months.

A large variety of outcome measures was applied to evaluate the Ginkgo effects. In all 8 studies in which a global assessment of the treatment effect was made, either by the patient, or the physician / caregiver, or both of them, Ginkgo treatment was superior to placebo. Statistical significance was reported in 4 studies.

Clinical rating scales were used in 8 trials as well. The application of the Crichton Geriatric Scale in its original or in a modified format yielded a significantly positive effect for Ginkgo in one trial²⁸, no effect in another trial⁹³, while in a third study no results for this rating scale were given⁸⁴. Administration of the Plutchic Geriatric Rating Scale yielded a beneficial Ginkgo effect in 2 trials^{85,86}. In the study by Halama et al both the Sandoz Clinical Assessment Geriatric Scale (SCAG) - being the primary outcome measure - and the Syndrom Kurz Test (SKT) revealed a

(statistically) significant effect of Ginkgo⁹³. Together with the SGRS, another clinical rating scale, the SCAG was also assessed in the trial reported by Pidoux et al, again with a positive result⁸⁸. A geriatric clinical evaluation scale (EACG), comparable with the SCAG, showed positive results²⁶, just as a self-rating clinical symptoms questionnaire^{91,92}. In the trial by Halama et al clinical symptom rating gave inconsistent results⁹³. In 10 trials ample attention was paid to psychometric testing of any *Ginkgo biloba* effect, through a wealth of different tests and test batteries (see table 6).

The influence of Ginkgo treatment on daily life activities was assessed in 3 trials^{81,83,85}, in each case with a positive result. Neurophysiological characteristics (EEG, ECG, CCG, saccadic eye movements, blood pressure, pulse rate, etc.), the presence and severity of physical signs and functional symptoms, state of anxiety^{91,92}, and clinical and laboratory examination results were used as other endpoints to quantify the treatment effect.

All trials but one⁹³ failed to explain the relative importance of each of the outcome measures, but almost all authors could report beneficial effects of Ginkgo treatment for (almost) all endpoints considered^{26,28,81-83,85,86,88,89,91,92,94,96,97}. Israel et al reported a positive Ginkgo effect on 3 out of 4 cognitive functions⁸⁴, Geßner et al reported a positive effect for global evaluation and for two reaction time tests, but not for several other psycho(physio)logical tests⁹⁰, and the positive results reported by Halama et al did not extend to the Crichton scale and the craniocorpographic measurements. When compared to reference drugs, Ginkgo did slightly better than vincamine⁸⁷, but slightly worse than dihydroergotoxin⁹⁵.

5 DISCUSSION

We have identified 55 randomized clinical trials of *Ginkgo biloba* and cerebral insufficiency. We have tried to categorize them under six different headings according to the target health problem addressed. Most of the studies were conducted in Germany and France: 31 and 16 trials, respectively. Occasionally, patients were recruited in other countries: Italy (3 trials), UK (2 trials), USA, Sweden, and Austria (1 trial each). In each category almost all trials which compared Ginkgo with placebo treatment, showed a statistically significant effect or a positive trend in favour of Ginkgo. Summarizing and pooling of these results appeared to be almost unfeasible, not only because of the heterogeneity of study design aspects (in particular outcome measures and the success criteria based on them), but also because of the large variety in reporting style. In many articles and study reports results were presented for only part of the outcome results. Statistical test results were not recorded in a systematical way either. This kind of problems is reflected in the contents of the summary tables. A serious shortcoming of many studies is that hardly any attention is paid to the hierarchy of the outcome measures and to the clinical meaning of the recorded effects. Actually, this is why we decided to present the study results in a rather detailed fashion, instead of trying to reduce the outcome information to a more aggregate and summary level. Major methodological weaknesses of many trials were, in addition to the outcome definition, the choice and description of the patient definition and eligibility criteria, and the small sample sizes. We would like to emphasize that, generally speaking, the scientific information contained by the more recent trials, especially some trials which focussed on dementia, should be valued considerably higher than the information that we got from earlier trials.

Let us now return to the key question. Can Ginkgo treatment (120-240 mg per day, during 3 or more months) be regarded effective in patients with dementia, cognitive impairment, and related disorders? Or, stated differently, will the aggregate clinical study results encourage (relatives of) patients to ask for Ginkgo preparations, physicians to prescribe them, and health authorities to consider licensing Ginkgo-based drugs?

The evidence rooted in the controlled trials in this overview is rather convincing. In general, the results point to a significant effect of *Ginkgo biloba* treatment. The results of the studies which addressed the same target health problem and therefore shared the same category, show a remarkable consistency, although in several trials the positive effect did not concern all outcome measures. It should be stressed, however, that only a minority of the trials can stand the test of criticism posed by the currently accepted guidelines and prescriptions regarding the methodology of clinical intervention trials in general and for dementia in particular. As a consequence, a substantial part of the available evidence is contributed by a rather small proportion of the clinical studies, most of which have been published quite recently. Although the time-related difference in design choices may distort the image, we cannot get away from the impression that, generally speaking, the effect sizes reported for the older trials exceed those reported for the newer trials. On average, these had larger sample sizes, providing for sufficient statistical power to make even rather non-impressive treatment effects statistically significant.

In our opinion, the clinical importance of the effects showed by the latest, high-quality studies remains debatable. Such a debate is seriously hampered by the high level of abstraction of the concepts underlying many of the questionnaires, tests and rating scales which tend to be used as outcome meas-

ures. The outcome of such a debate will be partly influenced by one's ideas regarding the treatability and reversibility of the health problem. It goes without saying that these conclusions and considerations rest upon the assumption that the picture of the efficacy of *Ginkgo* which emerges from the scientific literature, is not seriously biased by any form of selective reporting and publication.

REFERENCES

1. Van Dongen M, Van Rossum E, Knipschild P. Efficacy of *Ginkgo biloba* extracts: evidence from randomized clinical trials. In: Van Beek TA, ed. *Ginkgo biloba*. Amsterdam: Harwood Academic Publishers, 1999 (*In press*).
2. Fünfgeld EW, ed. *Rökan (Ginkgo biloba)*. Recent results in pharmacology and clinic. New York: Springer-Verlag, 1988.
3. DeFeudis FV. *Ginkgo biloba extract (EGb 761)*. Pharmacological activities and clinical applications. New York: Elsevier, 1991.
4. Oberpichler Schwenk H, Krieglstein J. Pharmakologische Wirkungen von *Ginkgo biloba* Extrakt und Inhaltsstoffen. *Pharm Unserer Zeit* 1992; 21: 224-35.
5. Hitzengerger G. Zur Wirkung von *Ginkgo biloba* Spezialextrakt (EGb 761, Tebofortan). *Wien Med Wochenschr* 1992; 142: 371-9.
6. DeFeudis FV. *Ginkgo biloba extract (EGb 761)*. From chemistry to the clinic. Wiesbaden: Ullstein Medical, 1998.
7. Sticher O. *Ginkgo biloba* Analytik und Zubereitungsformen. *Pharm Unserer Zeit* 1992; 21: 253-65.
8. Cooper H, Hedges LV, eds. *The handbook of research synthesis*. New York: Russell Sage Foundation, 1994.
9. Petitti DB. *Meta-analysis, decision analysis and cost-effectiveness. Methods for quantitative synthesis in medicine*. New York: Oxford University Press, 1994.
10. Chalmers I, Altman DG, eds. *Systematic reviews*. London: BMJ Publishing Group, 1995.
11. Warburton DM. Psycho-pharmacologie clinique de l'extrait de *Ginkgo biloba*. *Presse Med* 1986;15: 1595-604.
12. Weiß H, Kallischnigg G. *Ginkgo-biloba-Extrakt (EGb 761)*. Meta-Analyse von Studien zum Nachweis der therapeutischen Wirksamkeit bei Hirnleistungsstörungen bzw. arterieller Verschlusskrankheit. *Münch Med Wochenschr* 1991; 133: 138-42.
13. Herrschaft H. Zur klinischen Anwendung von *Ginkgo biloba* bei dementiellen Syndromen (Hirnleistungsstörungen bei vasculärer oder degenerativer ZNS Erkrankung). *Pharm Unserer Zeit* 1992; 21: 266-75.
14. Letzel H, Schoop W. *Ginkgo biloba* Extrakt EGb 761 und Pentoxifyllin bei Claudicatio intermittens. Sekundäranalyse zur klinischen Wirksamkeit. *Vasa* 1992; 21: 403-10.
15. Schulz V, Hänsel R. *Rationale Phytotherapie. Ratgeber für die ärztliche Praxis*. Ch.2: Zentrales Nervensystem. § 2.1. *Ginkgo* bei Hirnleistungsstörungen. Berlin: Springer, 1996.
16. Letzel H, Haan J, Feil WB. Nootropics: efficacy and tolerability of products from three active substance classes. *J Drug Dev Clin Pract* 1996; 8: 77-94.
17. Kanowski S. *Ginkgo-biloba-Spezialextrakt*. Nachgewiesene Wirksamkeit im Indikationsbereich Demenz. *Münch Med Wochenschr* 1997; 139: 39-42.
18. Schulz V, Hubner WD, Ploch M. Clinical trials with phyto-psychopharmacological agents. *Phytomedicine* 1997; 4: 379-87.
19. Kleijnen J, Knipschild P. *Ginkgo biloba* for intermittent claudication and cerebral insufficiency. In: Kleijnen J, ed. *Food supplements and their efficacy*. Thesis. Maastricht: Rijksuniversiteit Limburg, 1991: 83-94.
20. Kleijnen J, Knipschild P. *Ginkgo biloba* for cerebral insufficiency. *Br J Clin Pharmacol* 1992; 34: 352-8.
21. Kleijnen J, Knipschild P. *Ginkgo biloba*. *Lancet* 1992; 340: 1136-9.
22. Hopfenmüller W. Nachweis der therapeutischen Wirksamkeit eines *Ginkgo biloba* Spezialextrak-

- tes. Meta-Analyse von 11 klinischen Studien bei Patienten mit Hirnleistungsstörungen im Alter. *Arzneimittelforschung* 1994; 44: 1005-13.
23. Schmidt U, Rabinovici K, Lande S. Einfluß eines Ginkgo-Spezial-Extraktes auf die Befindlichkeit bei zerebraler Insuffizienz. *Münch Med Wochenschr* 1991; 133 (Suppl. 30): S15 - S18.
 24. Brüchert E, Heinrich SE, Ruf-Kohler P. Wirksamkeit von LI 1370 bei älteren Patienten mit Hirnleistungsschwäche. Multizentrische Doppelblindstudie des Fachverbandes Deutscher Allgemeinärzte. *Münch Med Wochenschr* 1991; 133 (Suppl. 30): S9 - S14.
 25. Meyer B. Étude multicentrique randomisée à double insu face au placebo du traitement des acouphènes par l'extrait de Ginkgo biloba. *Presse Med* 1986; 15: 1562-4.
 26. Taillandier J, Ammar A, Rabourdin JP, et al. Traitement des troubles du vieillissement cérébral par l'extrait de Ginkgo biloba. Étude longitudinale multicentrique à double insu face au placebo. *Presse Med* 1986; 15: 1583-7.
 27. Haguenaer JP, Cantenot F, Koskas H, Pierart H. Traitement des troubles de l'équilibre par l'extrait de Ginkgo biloba. Étude multicentrique à double insu face au placebo. *Presse Med* 1986; 15: 1569-72.
 28. Vorberg G, Schenk N, Schmidt U. Wirksamkeit eines neuen Ginkgo-biloba-Extraktes bei 100 Patienten mit zerebraler Insuffizienz. *Herz + Gefäße* 1989; 9: 936-41.
 29. Eckmann F. Hirnleistungsstörungen - Behandlung mit Ginkgo-biloba-Extrakt. Zeitpunkt des Wirkungseintritts in einer Doppelblindstudie mit 60 stationären Patienten. *Fortschr Med* 1990; 108: 557-60.
 30. Wesnes K, Simmons D, Rook M, Simpson P. A double-blind placebo-controlled trial of Tanakan in the treatment of idiopathic cognitive impairment in the elderly. *Human Psychopharmacology* 1987; 2: 159-69.
 31. Hartmann A, Frick M. Wirkung eines Ginkgo-Spezial-Extraktes auf psychometrische Parameter bei Patienten mit vaskulär bedingter Demenz. *Münch Med Wochenschr* 1991; 133 (Suppl. 30): S23 - S25.
 32. ICH3. *Structure and contents of clinical study reports. ICH-Efficacy Topic 3.* The Third International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use. Yokohama, 1995.
 33. McKhann G, Drachman D, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984; 34: 939-944.
 34. APA. *DSM-III-R. Diagnostic and Statistical Manual of Mental Disorders.* 3th ed. Washington, D.C.: American Psychiatric Association, 1987.
 35. WHO. *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines.* Geneva: World Health Organisation, 1992.
 36. APA. *Diagnostic and statistical manual of mental disorders: DSM-IV.* Prepared by Task Force on DSM-IV. 4th ed. Washington, D.C.: American Psychiatric Association, 1994.
 37. CBO. *Herziening consensus diagnostiek bij het dementiesyndroom.* Utrecht: Centraal Begeleidingsorgaan voor de Intercollegiale Toetsing, 1997.
 38. Small GW, Rabins PV, Barry PP, et al. Diagnosis and treatment of Alzheimer Disease and related disorders. Consensus statement of the American Association for Geriatric Psychiatry, the Alzheimer's Association, and the American Geriatrics Society. *JAMA* 1997; 278: 1363-71.
 39. Crook TH, Bartus RT, Ferris SH, Whitehouse P, Cohen GD, Gershon S. Age-associated memory impairment: proposed diagnostic criteria and measure of clinical change - report of a National Institute of Mental Health working group. *Dev Neuropsych* 1986; 2: 261-76.
 40. Crook TH, Larrabee GJ, Youngjohn JR. Diagnosis and assessment of age-associated memory impairment. *Clin Neuropharmacol* 1990; 13: S81-91.
 41. Amaducci L, Angst J, Bech P, Benkert O, Bruinvels J, Engel RR. Consensus conference on the methodology of clinical trials of 'nootropics', Mu-

- nich, June 1989. Report of the Consensus Committee. *Pharmacopsychiatry* 1990; 23: 171-5.
42. Bundesgesundheitsamt. Empfehlungen zum Wirksamkeitsnachweis von Nootropika im Indikationsbereich "Demenz" (Phase III). *Bundesgesundheitsblatt* 1991; 7: 342-50.
 43. Menges K. Proof of efficacy of nootropics for the indication "Dementia" (Phase III) - Recommendations. *Pharmacopsychiatry* 1992; 25: 126-35.
 44. CPMP. *Note for guidance: antidementia medicinal products*. CPMP Working Party on Efficacy of Medicinal Products. III/3705-91-EN, Draft 5. Brussels: Commission of the European Communities, 1992.
 45. CPMP. *Note for guidance on medicinal products in the treatment of Alzheimer's disease*. London: The European Agency for the Evaluation of Medicinal Products, Human Medicines Evaluation Unit, Committee for Proprietary Medicinal Products, 1997.
 46. Hofferberth B. The efficacy of EGb 761 in patients with senile dementia of the Alzheimer type, a double-blind, placebo-controlled study on different levels of investigation. *Human Psychopharmacology* 1994; 9: 215-22.
 47. 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.
 48. Haase J, Halama P, Hörr R. Wirksamkeit kurzdauernder Infusionsbehandlungen mit Ginkgo-biloba-Spezialextrakt EGb 761 bei Demenz vom vaskulären und Alzheimer-Typ. *Z Gerontol Geriatr* 1996; 29: 302-9.
 49. Maurer K, Ihl R, Dierks T, Frölich L. Clinical efficacy of Ginkgo biloba special extract EGb 761 in dementia of the Alzheimer type. *J Psychiatr Res* 1997; 31: 645-55.
 50. Le Bars PL, Katz MM, Berman N, et al. A placebo-controlled, double-blind, randomized trial of an extract of Ginkgo biloba for dementia. *JAMA* 1997; 278: 1327-32.
 51. Ihl R, Dierks T, Martin E, Frölich L, Maurer K. Ginkgo biloba Extrakt EGb 761 bei der Demenz vom Alzheimer Typ. Abstract. 15th Danube Symposium of Psychiatry, Regensburg, 1992.
 52. Weitbrecht WV, Jansen W. Doubleblind and comparative (Ginkgo biloba versus placebo) therapeutic study in geriatric patients with primary degenerative dementia - a preliminary evaluation. In: Agnoli J, Rapin R, Scapagnini V, Weitbrecht WV, eds. *Effects of Ginkgo Biloba Extract on organic cerebral impairment*. Montrouge: John Libbey Eurotext Ltd., 1985: 91-99.
 53. Weitbrecht W-U, Jansen W. Primär degenerative Demenz: Therapie mit Ginkgo-biloba-Extrakt. Placebo kontrollierte Doppelblind- und Vergleichsstudie. *Fortschr Med* 1986; 104: 199-202.
 54. Halama P. Ginkgo biloba. Wirksamkeit eines Spezialextrakts bei Patienten mit zerebraler Insuffizienz. *Münch Med Wochenschr* 1991; 133: 190-4.
 55. Halama P. Befindlichkeitsbeurteilung und Psychometrie. Testung der Ginkgo-biloba-Wirkung bei Patienten einer Neurologischen Fachpraxis. *Münch Med Wochenschr* 1991; 133 (Suppl. 30): S19 - S22.
 56. Vesper J, Hänsgen K-D. Efficacy of Ginkgo biloba in 90 outpatients with cerebral insufficiency caused by old age. Results of a placebo-controlled double-blind trial. *Phytomedicine* 1994; 1: 9-16.
 57. Chartres J-P, Bonnan P, Martin G. Réduction de posologie de médicaments psychotropes chez des personnes âgées vivant en institution. Étude à double-insu chez des patients prenant soit de l'extrait de Ginkgo biloba 761 soit du placebo. *Psychologie Médicale* 1987; 19: 1365-75.
 58. Israël L, Dell'Accio E, Martin G, Hugonot R. Extrait de Ginkgo biloba et exercices d'entraînement de la mémoire. Evaluation comparative chez des personnes âgées ambulatoires. *Psychologie Médicale* 1987; 19: 1431-9.
 59. Rai GS, Shovlin C, Wesnes KA. A double-blind, placebo controlled study of Ginkgo biloba extract ('tanakan') in elderly outpatients with mild to moderate memory impairment. *Curr Med Res Opin* 1991; 12: 350-5.
 60. Franco L, Cuny G, Nancy FMC. Étude multicentrique de l'efficacité de l'extrait de Ginkgo biloba

- (EGb 761) dans le traitement des troubles mnésiques liés à l'âge. *La Revue de Gériatrie* 1991; 16: 191-5.
61. Allain H, Raoul P, Lieury A, LeCoz F, Gandon JM, d'Arbigny P. Effect of two doses of ginkgo biloba extract (EGb 761) on the dual-coding test in elderly subjects. *Clin Ther* 1993; 15: 549-58.
 62. Semlitsch HV, Anderer P, Saletu B, Binder GA, Decker KA. Cognitive psychophysiology in nootropic drug research: effects of Ginkgo biloba on event-related potentials (P300) in age-associated memory impairment. *Pharmacopsychiatry* 1995; 28: 134-42.
 63. Oswald WD, Hörr R, Oswald B, Steger W, Sappa J. Zur Verbesserung fluidier, kognitiver Leistungen mit Ginkgo-biloba-Spezialextrakt EGb 761[®] bei Patienten mit leichten bis mittelschweren Hirnleistungsstörungen im Alter. *Zeitschrift für Gerontopsychologie und -psychiatrie* 1997; 10: 133-46.
 64. Crook TH. Assessment of drug efficacy in age-associated memory impairment. *Adv Neurol* 1990; 51: 211-6.
 65. Chesseboeuf L, Herard J, Trevin J. Étude comparative de deux vasorégulateurs dans les hypacusies et les syndromes vertigineux. *Médecine du Nord et de l'Est*. 1979; 3: 534-9.
 66. Natali R, Rachinel J, Pouyat PM. Essai comparatif croisé en O.R.L. de deux médicaments vaso-actives. *Cahiers d' Oto Rhino Laryngologie* 1979; 14: 185-90.
 67. Schwerdtfeger F. Elektronystagmographisch und klinisch dokumentierte Therapieerfahrungen mit rōkan[®] bei Schwindelsymptomatik. *Therapiewoche* 1981; 31: 8658-67.
 68. Claussen C-F. Mit Ginkgo biloba wird Ihr Patient wieder standfest. Craniocorpographie zeigt statistisch signifikanten Rückgang der Vertigo- und Ataxie-Symptomatik im Doppelblindversuch. *Ärztliche Praxis* 1984; 36: 3-8.
 69. Claussen C-F, Kirtane MV. Randomisierte Doppelblindstudie zur Wirkung von Extractum Ginkgo biloba bei Schwindel und Gangunsicherheit des älteren Menschen. In: Claussen C-F, ed. *Presbyvertigo, Presbyataxie, Presbytinnitus: Gleichgewichts- und Sinnesstörungen im Alter*. Berlin: Springer Verlag, 1985: 103-15.
 70. Hamann K-F. Physikalische Therapie des vestibulären Schwindels in Verbindung mit Ginkgo-biloba-Extrakt. Eine posturographische Studie. *Therapiewoche* 1985; 35: 4586-90.
 71. Dubreuil C. Essai thérapeutique dans les surdités cochléaires aiguës. Étude comparative de l'extrait de Ginkgo biloba et de la nicergoline. *Presse Med* 1986; 15: 1559-61.
 72. Meyer B. Étude multicentrique des acouphènes. Epidémiologie et thérapeutique. *Ann Otolaryngol Chir Cervicofac* 1986; 103: 185-8.
 73. Holgers KM, Axelsson A, Pringle I. Ginkgo biloba extract for the treatment of tinnitus. *Audiology* 1994; 33: 85-92.
 74. Hoffmann F, Beck C, Schutz A, Offermann P. Ginkgoextrakt EGb 761 (Tebonin)/HAES versus Naftidrofuryl (Dusodril)/HAES. Eine randomisierte Studie zur Horsturztherapie. *Laryngorhinootologie* 1994; 73: 149-52.
 75. Morgenstern C, Biermann E. Ginkgo-Spezialextrakt EGb 761 in der Behandlung des Tinnitus aurium. Ergebnisse einer randomisierten, doppelblinden, plazebokontrollierten Studie. *Fortschr Med* 1997; 115: 7-11.
 76. Eckmann F, Schlag H. Kontrollierte Doppelblind Studie zum Wirksamkeitsnachweis von Tebonin forte bei Patienten mit zerebrovaskulärer Insuffizienz. *Fortschr Med* 1982; 100: 1474-8.
 77. Hofferberth B. Ginkgo-biloba-Spezialextrakt bei Patienten mit hinorganischem Psychosyndrom. Prüfung der Wirksamkeit mit neurophysiologischen und psychometrischen Methoden. *Münch Med Wochenschr* 1991; 133 (Suppl. 30): S30 - S33.
 78. Halama P. Was leistet der Spezialextrakt EGb 761? Therapie mit Ginkgo biloba bei Patienten mit cerebrovaskulärer Insuffizienz und therapieresistenter depressiver Symptomatik. Ergebnisse einer plazebokontrollierten, randomisierten doppelblinden Pilotstudie. *Therapiewoche* 1990; 40: 3760-5.
 79. Stocksmeier U, Eberlein M. Depressive Verstimmung bei Hirnleistungsstörungen. Wirkung eines Ginkgo-biloba-Extraktes in Doppelblind-Studie überprüft. *TW Neurologie Psychiatrie* 1992; 6: 74-76.

80. Schubert H, Halama P. Primär therapieresistente depressive Verstimmung älterer Patienten mit Hirnleistungsstörungen: Wirksamkeit der Kombination von Ginkgo-biloba-Extrakt EGb 761 mit Antidepressiva. *Geriatric Research* 1993; 3: 45-53.
81. Moreau P. Un nouveau stimulant circulatoire cérébral. *La Nouvelle Presse Médicale* 1975; 4: 2401-2.
82. Bono Y, Mouren P. L'insuffisance circulatoire cérébrale et son traitement par l'extrait de Ginkgo biloba. *Méditerranée Médicale* 1975; 3: 59-62.
83. Augustin P. Le Tanakan en gériatrie. Étude clinique et psychométrique chez 189 malades d'hospice. *Psychologie Médicale* 1976; 8: 123-30.
84. Israël L, Ohlman T, Delomier Y, Hugonot R. Étude psychométrique de l'activité d'un extrait végétal au cours des états d'involution sénile. *Lyon Méditerranée Médicale* 1977; 8: 1197-9.
85. Agnoli A. Relazione clinica sulla specialità "Tebonin forte". L'Aquila: Istituto Universitario di Medicina e Chirurgia L'Aquila, Clinica Neurologica, 1980.
86. Dieli G, La Mantia V, Sietta M, Costanzo E. Essai clinique à double insu du tanakan dans l'insuffisance cérébrale chronique. *Lavoro Neuropsychiatrico* 1981; 68: 1-7.
87. Haan J, Reckermann U, Welter FL, Sabin G, Müller E. Ginkgo-biloba-Flavonglykoside. Therapiemöglichkeit der zerebraler Insuffizienz. *Med Welt* 1982; 33: 1001-5.
88. Pidoux B, Bastien C, Niddam S. Clinical and quantitative EEG double-blind study of Ginkgo biloba extract (GBE). *J Cereb Blood Flow Metab* 1983; 3 (Suppl. 1): S556 - S557.
89. Teigeler R, Pieprzyk L. Ginkgo biloba bei zerebraler Insuffizienz. *Ärztliche Praxis* 1984; 36: 1374.
90. Geßner B, Voelp A, Klasser M. Study of the long-term action of a Ginkgo biloba extract on vigilance and mental performance as determined by means of quantitative pharmaco-EEG and psychometric measurements. *Arzneimittelforschung* 1985; 35: 1459-65.
91. Arrigo A, Cattaneo S. Clinical and psychometric evaluation of Ginkgo biloba extract in chronic cerebro-vascular diseases. In: Agnoli J, Rapin R, Scapagnini V, Weitbrecht WV, eds. Effects of Ginkgo Biloba Extract on organic cerebral impairment. Montrouge: John Libbey Eurotext Ltd., 1985: 85-89.
92. Arrigo A. Behandlung der chronischen zerebrovaskulären Insuffizienz mit Ginkgo-biloba-Extrakt. *Therapiewoche* 1986; 36: 5208-18.
93. Halama P, Bartsch G, Meng G. Hirnleistungsstörungen vaskulärer Genese. Randomisierte Doppelblindstudie zur Wirksamkeit von Ginkgo biloba Extrakt. *Fortschr Med* 1988; 106: 408-12.
94. Hofferberth B. Einfluß von Ginkgo biloba Extrakt auf neurophysiologische und psychometrische Meßergebnisse bei Patienten mit hirnganischem Psychosyndrom. Eine Doppelblindstudie gegen Placebo. *Arzneimittelforschung* 1989; 39: 918-22.
95. Gerhardt G, Rogalla K, Jaeger J. Medikamentöse Therapie von Hirnleistungsstörungen. Randomisierte Vergleichsstudie mit Dihydroergotoxin und Ginkgo biloba Extrakt. *Fortschr Med* 1990; 108: 384-8.
96. Maier-Hauff K. LI 1370 nach zerebraler Aneurysma-Operation. Wirksamkeit bei ambulanten Patienten mit Störungen der Hirnleistungsfähigkeit. *Münch Med Wochenschr* 1991; 133 (Suppl. 30): S34 -S37.
97. Gräßel E. Einfluß von Ginkgo biloba Extrakt auf die geistige Leistungsfähigkeit. Doppelblindstudie unter computerisierten Meßbedingungen bei Patienten mit Zerebralinsuffizienz. *Fortschr Med* 1992; 110: 73-6.

Table 1. RCTs on the efficacy of *Ginkgo biloba* special extract in patients suffering from cerebral insufficiency: dementia (AD, VD, mixed type).

Author, year	Study population			Intervention
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Weitbrecht and Jansen, 1985, 1986	dementia: primary degenerative (mild, moderate): HIS ≤ 7 no further specifications	40 20 / 20 38 18 / 20	60-80 (71.7) 14 / 26	Tebonin® forte oral, solution 120 mg/day 3 m.
Halama, 1991a, 1991b	cerebral insufficiency: degenerative + vascular dementia: 1. single senile (ICD 290.0) 2. presenile (ICD 290.1) 3. arteriosclerotic (ICD 290.4) no formal criteria stated MWT: premorbid intelligence	50 42 20 / 22	35-85 15 / 37	Kaven® (LI 1370) oral, tablets 150 mg/day 12 w.
Hartmann and Frick, 1991	vascular dementia: CT-scan typical symptoms HOPS: ≥ 1 symptom within each of 4 symptom groups present	52 26 / 26 48	62.5 34 / 18	Kaven® (LI 1370) oral, solution (drink) 150 mg/day 3 m.

Outcome measure	Result	Remarks
clinical rating scales:		parallel
- general feeling (self-assessment by patient) (1=excellent; 2=good; 3=poor; 4=bad)		vs placebo source: n.d.
- general health (physician)	+0.9 vs +0.0 points	also comparison
- Crichton Geriatric Scale; 11 items (1-5): 11-55	+5 vs +0 points	with ergot alkaloids (n=20)
- SCAG; 19 items (1-7): 11-133	+23 vs +2 points	pretreatment duration of symptoms: 8.8 yrs.
psychometric tests:		
- figure-symbol test (Wechsler)	+2.5 vs -1.0 points	
- number-repeat test (digit span) (Wechsler)	+2.0 vs +0.0 points	
- flicker fusion frequency	+1.8 vs -0.4 Hz	
- reaction time	+0.33 vs +0.02 sec	
Syndrom Kurz Test (Q-27)	+4.15 vs +1.32 (p<0.001)	parallel
global assessment of efficacy (patient)	13 / 4 / 3 vs 4 / 3 / 15 (p<0.001)	vs placebo
global assessment of efficacy (physician): (‘good’ / ‘moderate’ / ‘unsatisfactory’)	12 / 4 / 4 vs 2 / 3 / 17 (p<0.001)	ambulatory patients (1 neurological practice)
11 subjective clinical CI-symptoms: memory, forgetfulness, concentration, fatigue, performance, motivation, depression, anxiety, dizziness, headache, tinnitus; 4-point scale: severe, moderate, mild, absent	n. improved, Ginkgo vs. placebo: p <0.005: 2 symptoms p <0.05: 5 symptoms n.s., tendency: 1 symptom n.s.: 3 symptoms (anxiety, memory gap, concentration)	dropouts: 5 vs 3 baseline SKT score 5.25 vs 6.64
general fluent intelligence:		
number connection test (NAI-ZVT)	+7.2 sec. vs +0.0 sec (n.s.)	
blood pressure, pulse rate		
global assessment of efficacy (patient)		parallel
global assessment of efficacy (physician) (‘good’ / ‘moderate’ / ‘unsatisfactory’)	13 vs 8 ‘good’, 7 vs 10 ‘moderate’, 2 vs 5 ‘poor’ (n.s.)	vs placebo
12 clinical CI-symptoms (hetero-assessment) memory, forgetfulness, concentration, fatigue, performance, motivation, depression, anxiety, dizziness, headache, tinnitus, confusion; 4-point scale: severe, moderate, mild, absent	3 severity categories improvement in both groups difference Ginkgo-placebo: n.s.	out-patients: neurological dept. of university hospital exclusion criteria dropouts: 2 vs 2
psychometric testing:		
- Grünberger Verbale Gedächtnis test	+2.55 vs +0.66 points (p<0.05)	
- trail making test (ZVT)	+8 sec vs +7 sec (n.s.)	



Author, year	Study population			Intervention
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Vesper and Hänsgen, 1994	cerebral insufficiency: 1. MMS (Folstein) 2. CI-scale: >19 3. MWT-B \geq 80 4. DPD >2 5. K-HIS 6. medical examination	90 45 / 45 86 42 / 44	55-80	Kavenin [®] (LI 1370) oral, tablets 150 mg/day 12 w.
Hofferberth, 1994	dementia: Alzheimer type: - clinical diagn., incipient - Blessed Dementia Scale: sum part A: 0 - 16 sum part B: 9.5 - 30.5 - HIS: <4 - CT-scan	40 21 / 19	50-75 28 / 14	Tebonin [®] forte oral, tablets 80 mg/day 3 m.
Kanowski <i>et al.</i> , 1996	dementia: AD or MID: - clinical diagnosis (physician); according to DSM-III-R mild - moderate - CCT - Ischemic Score (Rosen) - MMSE: 13-25 - SKT: 6-18	216 156 79 / 77	70.1 51 / 105 MID / AD = 12 / 31	EGb 761 oral, capsules 240 mg/day 24 weeks

Outcome measure	Result	Remarks
CI-scale (subjective cerebral insuff. troubles; scale: improved, unchanged, deteriorated) rating of changes of 8 mental + physical functions: scale 1-7 (4 = no change) general intelligence test: KAI (information processing) computer-aided psychometric tests:	w. 1-6: 50% vs 32% improved w. 6-12: 53% vs 37% improved 4.7 - 5.0 vs 3.9 - 4.2 ($p < 0.05$) n.s. verum group sign. more improved +6 vs +2 ms/10 ($p < 0.05$) 0 vs -2 ($p < 0.05$) +38% vs +14% correct solutions verum group sign. more improved	parallel vs placebo out-patients: 6 neurologist / psy- chiatrist practices 4 GPs 1 internal specialist
<i>primary outcome measure:</i> - Syndrom Kurz Test (SKT): memory + attention, sum score 9 subtests (Q-27)	+5 vs -2 ($p < 0.001$) 50.4% vs 0.0% > +5 points	parallel vs placebo hospitalized patients excluded: MID,
<i>secondary outcome measures:</i> - SCAG: change in psychopathology 18 items (scale 1-7), sum score: 18-126 global impression - Vienna Determination test (reaction test) no. correct on steps 10, 12, 15 (max.: 180) - saccade test (duration, latency + maximum rate rapid eye movements (electrooculography)) - EEG-parameters: theta/alpha quotient - global assessment efficacy (physician): 'good', 'satisfactory', 'not satisfactory'	+29 vs +0 verum better on most items verum better ($p < 0.05$) step 10: +22 vs +4; step 12: +12 vs +2; step 15: +9 vs +2 latency: +65 vs -5 ms duration: +57 vs +2 ms +17% vs +0% ($p < 0.01$) 76% vs 0% 'good', 24% vs 42% 'satisfactory'	dementia with known cause, advanced dementia, depression, etc.
<i>primary outcome measures:</i> - clinical global impression of change (CGI-2) by physician; scale: 1-7 response criterion: '(very) much improved' - Syndrome Kurz Test (SKT): memory + attention, sum score 9 subtests (Q-27) response criterion: decrease \geq 4 points - Nuremberg Geriatric Observation Scale (NAI-NAB); behaviour: coping with everyday tasks, independence of care; relative-rating response criterion: decrease \geq 2 points - responder analysis: \geq 2 of 3 POMs	32% vs 17% +2.2 vs +0.8 38% vs 18% ($p < 0.05$) +0.9 vs +0.6 33% vs 23% ($p < 0.1$) 28% vs 10% ($p = 0.005$)	parallel vs placebo out-patients 41 centres: practices of GPs, neurologists, psychiatrists, inter- nists excluded: pronounced cerebral atrophy, full-time care unavoidable 4 w. run-in
<i>secondary outcome measures:</i> - Montgomery Asberg Depression Rating Scale - actometry (subsample: 38 vs 38) - EEG (subsample: 38 vs 38)	+3.3 vs +2.5	205: 1-to-T analysis effect in AD slightly larger than in MID

Author, year	Study population		Intervention	
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Haase <i>et al.</i> , 1996	dementia: AD, MID, mixed: - clinical diagnosis according to DSM-III-R - moderate - HIS (Rosen modification) - Global Deterioration Scale (GDS): 7 stages, stages 4-5 - MMST	40 20 / 20	45-87 16 / 24	EGb 761 i.v. (infusion) 200 mg/day 4 w. (4 days / week)
Le Bars <i>et al.</i> , 1997	dementia: AD, MID: - according to DSM-III-R, ICD-10 - mild to moderately severe: - MMSE: 9-26 - Global Deterioration Scale (GDS): 3-6	327 166 / 161 309 155 / 154 IIT-analysis 236 AD	45-90 143 / 166	EGb 761 oral, tablets 120 mg/day 52 w.
Maurer <i>et al.</i> , 1997	dementia: AD: - according to DSM-III-R, NINCDS-ADRDA - Brief Cognitive Rating Scale (BCRS): 3-5 - HIS ≤ 4 - CT-scan: diffuse atrophy or no pathological signs	20 10 / 10 18 9 / 9	50-80 9 / 9	Tebonin® forte oral, tablets 240 mg/day 3 m.

Outcome measure	Result	Remarks
<i>primary outcome measures:</i>		
1. Nuremberg Geriatric Observation Scale (NAI-NAB); behaviour: coping with everyday tasks, independence of care; relative-rating response criterion: decrease ≥ 2 points	+3.6 vs +0.3 points ($p < 0.01$) 17 vs 3 improved	parallel vs placebo out-patients, neuro-psychiatric practice prestratification:
2. clinical global impression of change (CGI-2) by physician; scale: 1-7	improved: 17 vs 7 ($p < 0.0001$) 1/2/3: 1/9/7 vs 0/0/7	- NAB-score - HIS;
3. KAI (information processing)	improved: 7 vs 1 ($p < 0.03$)	no interaction
<i>secondary outcome measures:</i>		
- Nuremberg Elderly Activities of Daily Life scale (NAI-NAA); 20 items; sum: 20-60	+8.0 vs +0.5 points ($p = 0.003$)	
- improvement of strain due to main symptom	forgetfulness: 17 vs 6 ($p = 0.002$)	
- Zung's Selfrating Depression Scale (SDS)	+6 vs +0 points ($p < 0.0001$)	
<i>primary outcome measures:</i>		
1. cognitive impairment:		parallel vs placebo
- ADAS-Cog (cognitive subscale) performance based (memory, language, praxis, orientation); 11 items; sum: 0-70	-0.1 vs -1.5 ($p = 0.04$) $\geq +2$: 50% vs 29%	out-patients, without other sign. conditions 6 centers
2. daily living + social behavior:		placebo run-in: 2 w. compliance:
- Geriatric Evaluation by Relative's Rating Instrument (GERRI); 49 items; rating by caregiver, after 14-day observ.	+0.06 vs -0.08 ($p = 0.004$) improved: 37% vs 23%	pill count, deviation <20%
- GERRI-cognitive (21 items)		78 / 59 completed the trial
- GERRI-social (18 items)		202 in per protocol analysis
- GERRI-mood (10 items) item, subscale, total scale score: 1-5		same results for the AD subgroup
3. general psychopathology:		
- Clinical Global Impression of Change (CGIC); clinician; ordinal scale: 1-7	4.2 vs 4.2 ($p = 0.77$)	
<i>secondary outcome measures:</i>		
not yet published		
<i>primary outcome measure:</i>		
- Syndrom Kurz Test (SKT): memory + attention, sum score 9 subtests (0-27)	+2.89 vs -0.78 ($p = 0.013$) (baseline: 19.67 vs 18.11)	parallel vs placebo out-patients
<i>secondary outcome measures:</i>		
- number connection (ZVT)	positive trend (n.s.)	run-in: 1 w. excluded: advanced dementia, dementia of other aetiology, etc.
- ADAS:		
- cognitive subscale	+0.88 vs -0.03 (n.s.)	
- non-cognitive subscale	+1.78 vs +1.45 (n.s.)	
- clinical global impression (CGI): 1 - 7	4.3 vs 5.3 ($p = 0.069$)	
- electrophysiological parameters (AEP-300, EEG-topography)	positive trend (n.s.)	
- patient termination record	8 / 9 vs 5 / 9 preferred to continue	

Table 2. RCTs on the efficacy of *Ginkgo biloba* special extract in patients suffering from cerebral insufficiency: cognitive impairment.

Author, year	Study population	Number randomized analyzed (EG/other)	Age (years)	Intervention
	Diagnostic category		Sex (M/F)	Brand name Administration route Dose Duration
Chartres <i>et al.</i> , 1987	mild cognitive impairment:	44	81.8	Tanakan®
	- MMS (Folstein): 20-26		1 : 3	n.d.
	- GDS (Reisberg): stages 1-4	42		n.d.
	psychotropic treatment (neuroleptics, anti-depressants, etc.) >3m.	n.s.		6 m.
Israel <i>et al.</i> , 1987	early cognitive problems:	80	56-83 (68.4)	Tanakan®
	- subjective complaints		26 / 54	oral, solution
	- MMS (Folstein): 20-26	75		160 mg/day (2x 2 ml)
	absence of depression:			3 m.
	- GDS (Yesavage): <18			
	examination by physician + psychologist			
Wesnes <i>et al.</i> , 1987	cognitive impairment:	58	62-85	Tanakan®
	- mild, impairment	28 / 30	34 / 20	oral, tablets
	- Crichton Geriatric Behavioural Scale:	54		120 mg/day
	≥14	27 / 27		12 w.
	not >3 for any item (mild impairment of everyday functioning)			

Outcome measure	Result	Remarks
global effect, VAS (0-100 mm) (physician)	47.6 vs 46.2 (n.s.)	parallel
global effect (patient)	0.8 vs 1.0 (trend; n.s.)	vs placebo
level of cognitive functioning:		residents old people's
- MMS	+0.5 vs +0.0	home (≥ 6 m.)
- Global Deterioration Scale (GDS)	+0.22 vs -0.04	duration of psycho-
- Plutchik Geriatric Rating Scale (GRS)	+0.7 vs -0.6	tropic treatment:
development of several clinical signs (tremor, fall, incontinence for urine, etc.)	no difference	54.7 m.
quantity of psychotropic drugs prescribed:	score: 24.7 (0 m.) \rightarrow 11.7 (6 m.)	
- neuroleptics:	+2.3 vs +0.9	
- antidepressants:	+1.7 vs +2.6	
- tranquilizers / hypnotics:	+8.8 vs +10.6	
global effect: drug + training (patient)	only ginkgo group: extra effect	parallel
global effect: drug - training (patient)	88% vs 23% satisfied ($p < 0.001$)	vs placebo
memory battery for elderly, ambulatory:	unique effect ginkgo, independent of training effect: 1. n.s.	out-patients: 3 sources
4 factors: 1. long-term memory. 2. short-term memory. 3. learning rate. 4. fluidity.	2. $p < 0.01$. 3. not. 4. $p < 0.005$	factorial design:
consequences of cognitive impairment for daily life: questionnaire (Israel) (physician)		4 groups: 2 with memory training (11 sessions)
intellectual dynamics scale (Israel) (physician)		
Geriatric Depression Scale (Yesavage) (patient)		
scale of daily life difficulties (MacNair & Kahn scale) (patient)		
overall assessment of efficacy (patient)	27 vs 25 'very good'	parallel
overall assessment of efficacy (physician)	27 vs 25 'very good'	vs placebo
behavioural rating scale: quality of life	no difference in daily activities, but showed more interest	non-institutionalized, GP-setting
cognitive testing:		exclusion criteria
- digit span (WAIS), forward + backward	+0.8 vs +0.4	run-in period: 1-3 w.
- Benton Visual Retention Test (visuo-spatial)	+0.3 vs -0.3	blocks of 6
- automated psychometric test battery:	1. +0.5 vs +1.0 2. number correct: +7.5 vs -6.4; time: +14 vs -35 msec	HIS: $51x \leq 4$,
1. immediate word recall, 2. number matching.	3. % detected: +12.3% vs +15.7%	so probably early
3. rapid visual information processing.	4. number correct: +0.8 vs +0.9; time: +24 vs +9 msec 5. +13.5 vs +10.2. 6. number correct: +0.0 vs +10.2; time: -441 vs -65 ms	stage of degenerative dementia
4. choice reaction time. 5. digit symbol substitution. 6. word recognition	effect size 4 accuracy tests: $p < 0.05$ effect size 4 speed tests: $p < 0.05$	
mood + alertness (11 VAS-scales)		
medical examination: ECG, haematology, biochemistry		

Author, year	Study population	Intervention		
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Rai <i>et al.</i> , 1991	memory impairment: 1. mild-moderate signs, organic origin, >3 m., NINCDS-ADRDA classification 2. clinical + medical assessment: confirmation of diagnosis	31 27 12 / 15	67-89 7 / 20	Tanakan® oral, tablets 120 mg/day 6 m.
Franco <i>et al.</i> , 1991	age-associated memory problems in daily life: - CDS-score: 40-120 - Zung-scale <50 (no severe depression)	97 78 38 / 40	60-85 n.d.	Tanakan® n.d. 120 mg/day 12 m.
Allain <i>et al.</i> , 1993	age-related memory impairment (slight degree): - MMSE: 25-28 - objective cognitive impairment (immediate recall of word lists)	18 6x 3 subj.	60-80 (69.3) n.d.	Tanakan® (IPSEN) n.d. 600 mg / 320 mg acute

Outcome measure	Result	Remarks
MMSE (Folstein) (0-30)	+0.18 vs -0.29	parallel
Kendrick battery: detecting dementia in elderly:		vs placebo
- digit copying task	+5.55 vs -2.43 (p=0.017)	out-patients:
- object learning task	+0.82 vs +2.43	hospital, dept. of
digit recall test: no. correct before error	-2.5 vs + 5.3 (p=0.032)	geriatric medicine
(computerized) total no. correct	-1.7 vs +4.1	
magnitude of error	+26.7 vs -34.2	
classification task: no. attempted	+0.00 vs -0.17	
(computerized) median reaction time	+171 vs -37 msec (p=0.026)	
latency auditory event related potential (P300)		
EEG-frequency analysis		
<i>primary outcome measures:</i>	+11.6 vs +7.7	parallel
- CDS-score (cognitive deficits self-assessment scale (Mac Nair)): troubles of attention, concentration, psychomotor coordination, orientation, STM, LTM		vs nicergolin
- Zung scale (stability of mood)		not-institutionalized:
	+1.5 vs +0.3 (n.s.)	city practice own GP, 12 centres
<i>secondary outcome measures:</i>	70.6 vs 58.3 mm	
- global assessment efficacy, VAS (patient)	72.8 vs 58.3 mm	
- global assessment efficacy, VAS (physician)	35 vs 28: continue (p=0.007)	
- preference for post-trial treatment		
- development of functional signs (1. motor, 2. neuro-sensoric, 3. psycho-affective)	3. +2.9 vs +1.9	
relative speed of information processing:	no. correctly recalled,	cross-over
dual-coding test (evaluates memory coding of verbal material (words) and images (drawings) in relation to variable presentation times)	EGb 600 vs EGb 320 vs placebo:	EGb 761 600 mg vs
3 different batteries of memory tests (5 series of 8 drawings + 5 series of 8 words; presentation times: 1920, 960, 480, 240, 120 msec)	- Drawings, 1920: 5.3 ; 5.5 ; 5.1	EGb 761 320 mg vs
testing 1 h after drug administration	- Drawings, 960: 4.2 ; 4.3 ; 3.9	placebo
	- Drawings, 480: 3.2 ; 3.0 ; 3.2	source: 1 centre
	- Words, 1920: 3.3 ; 3.3 ; 3.3	Latin square design
	- Words, 960: 3.0 ; 2.9 ; 3.5	wash-out: 6 d.
	- Words, 480: 2.9 ; 2.6 ; 2.8	acute effect measured
	Drawings / Words difference:	
	- breakpoint: 480 vs 480 vs 960	
	- dual coding: 960 vs 960 vs 1920	



Author, year	Study population			Intervention
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Semlitsch <i>et al.</i> , 1995	age-associated memory impairment (AAMI): - age >50 - subjective everyday memory problems - gradual course memory loss - Grünberger Verbal Memory Test <29 - BVRT - MMSE ≥24 (no dementia)	48 24 / 24 43 21 / 22	51-79 (62) 19 / 29	Ginkobene® oral dosage 120 mg/day 57 d.
Oswald <i>et al.</i> , 1997	organic brain syndrome: - clinical diagnosis according to DSM-III-R, ICD-10, 290.x - mild to moderate - SCAG: sum 18 items: 40-90 mean items 2,3,6,8 >3.5 sum items 1 - 8: >24 items 1 - 8: not 7	226 195 103 / 92	60-91 81 / 145	EGb 761® oral, tablets 120 mg/day 12 w.

Outcome measure	Result	Remarks
psychometric tests		parallel
EEG brain-mapping		vs placebo
cognitive information processing:	no consistent and unequivocal	source: n.d.
psychophysiological investigations: two-tone auditory oddball paradigm, registration of long-latency Event-Related Potentials (ERP); P300 latency may reflect stimulus-evaluation time	changes on N1, P2, N2, P300 amplitudes, nor N1, P2, N2 latencies; P300 latency shortened by 31 msec (acute), 38 msec (chronic), 32 msec (superimposed) for Ginkgo	
subjective well-being: EWL-K checklist (self-reporting thymopsychic scale; subscales: activation, deactivation, assertiveness, fatigue, dizziness, high spirits, agitation, sensitiveness, annoyance, anxiety, depression, dreaminess)	acute treatment (w.0, 3 h): no difference chronic treatment (w.8, 0 h): GBE: increase in deactivation, fatigue, dizziness, anxiety superimposed treatment (w.8, 3 h): GBE: increase in dizziness	
Nuremberg Gerontopsychological Inventory (NAI): 3 domains:		parallel
1. fluid cognitive performance (objective test):		vs placebo
- number - symbol test (ZS-G)	45% vs 38% improved (0.174)	out-patients, living
- labyrinth-test (LT-G)	65% vs 53% improved (0.047)	at home or in an old
- number - connection test (ZVT-G)	54% vs 53% improved (0.438)	people's home;
responder: ≥ 2 of 3 tests	59.2% vs 48.9% ($p = 0.074$)	2 GP-practices
2. well-being (self-assessment):		exclusion: specific
- self-assessment scale (NAS)	49% vs 47% improved (0.401)	causes of brain
- daily activities scale (NAA)	35% vs 38% improved (0.673)	syndrome
- self-estimation scale (NSL)	52% vs 54% improved (0.657)	run-in period: 2 w.
responder: ≥ 2 of 3 scales	46.6% vs 45.7% ($p=0.447$)	compliance (urine,
3. function level (hetero-assessment)		pill count) >90%
- elderly rating (NAR): 9 items	81% vs 74% improved (0.133)	
responder: no deterioration	80.6% vs 73.9% ($p=0.133$)	
efficacy: respondent in ≥ 2 of 3 domains	n.s.	

Table 3. RCTs on the efficacy of *Ginkgo biloba* special extract in patients suffering from cerebral insufficiency: non-cognitive signs and symptoms.

Author, year	Study population			Intervention
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Chesseboeuf <i>et al.</i> , 1979	hearing loss / vertigo / tinnitus: (vascular origin, or ageing of inner ear, or due to injury, infarction) diagnostic criteria: n.d.	60 30 / 30	51.6 35 / 25	Tanakan® oral, solution 3 ml/day 2 m.
Natali <i>et al.</i> , 1979	hearing loss / tinnitus + vertigo: (partly) of vascular origin) severity of hearing loss: 9x <30 db, 9x >30 db	20	60 10 / 10	Tanakan® oral, solution 3 ml/day 1-2 m.
Schwerdtfeger, 1981	dizziness symptoms: origin: 7 categories neurotologic examination pathological findings of electronystagmogram (butterfly or L-pattern)	50 25 / 25 42 23 / 19	17-63 (43) 17 / 33	rökan® oral, tablets 60 mg/day (3 dragees) 2 m.
Claussen, 1984 Claussen and Kirtane, 1985	vertigo / ataxia symptoms: strong, subjective dizziness symptoms	50 25 / 25 33 14 / 19	50-71 (55.5) 17/16	rökan® oral, tablets 120 mg / day 12 w.
Hamann, 1985	vertigo: dizziness complaints due to Menière disease, vesti- bular neuropathy, trauma ≥1 out of 3 signs: dizzi- ness, visual motor changes (electronystagmography), body balance disturbances	40 20 / 20 35 17 / 18	46.2 13 / 27	Tebonin® forte oral, solution (drops) 160 mg/day 4 w. adjuvans therapy (supplementary to physical, vestibular training; 10 sessions)

Outcome measure	Result	Remarks
global effect: very good / good / moderate / no functional signs + symptoms:	5/13/8/4 vs 0/12/14/4	parallel vs nicergolin (NCG) source: n.d.
- hearing loss:	+0.46 vs +0.46 (n.s.)	
very good / good / moderate / no change	1/5/4/16 vs 2/1/6/15	
- tinnitus:	+1.13 vs +1.07 (n.s.)	
very good / good / moderate / no change	2/2/7/4 vs 3/4/6/1	
- vertigo:	+1.45 vs +0.83 (p<0.1)	
very good / good / moderate / no change	4/6/7/4 vs 1/3/6/5	
- audiogram:	+0.35 vs +0.29 (n.s.)	
% improved	33% vs 28%	
- electronystagmogram:	+0.8 vs +0.6 (n.s.)	
% improved	35% vs 13%	
global effect: (very) good / insufficient / no functional signs (presence, severity):	12/4/4 vs 1/12/7	cross-over vs cinnarizin (CIN)
- labyrinth examination	5/18 vs 1/18 '(very) good'	hospital patients
- audiometric balance	tinnitus:	only partly randomized
- electronystagmogram (ENG)	10/13 vs 0/13 '(very) good'	(until 13th patient):
	vertigo:	13 started with CIN
	2/2 vs 0/2 '(very) good'	wash-out: 1 m.
questionnaire	overall evaluation: number	parallel
electronystagmography:	improved / unchanged / impaired:	vs placebo
1. calorimetric (butterfly pattern, Claussen)	11/9/3 vs 3/4/12	source: n.d.
2. rotary chair (L-pattern, Claussen)	0 vs 7 relapsed after successful open trial	first open trial: 2 m.
neuro(oto)logical tests (10 parameters)		(rökan, 160 mg/day)
tonal audiogram (+ vocal, + impedance)		
questionnaire on subjective dizziness symptoms	+50% vs +20% improvement	parallel
lateral body sway: craniocorography (CCG)	+9.2 cm vs +2.7 cm (p<0.005)	vs placebo source: n.d.
vertigo questionnaire:	% cured or improved:	parallel
subjective dizziness complaints	ca. 75% vs ca. 75% (n.s.)	vs placebo
	% unchanged: 20% vs 18%	source: n.d.
Romberg's test:		no response to previous, 'classic' therapy
posturographic body sway:		
- anterior-posterior	+23.1 mm vs +15.5 mm	
- lateral	+21.2 mm vs +10.8 mm	
	ca. +41% vs ca. +26%	

Author, year	Study population			Intervention
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Dubreuil, 1986	cochlear deafness:	20	49.6	Tanakan®
	partly or total, ≥1 week	10 / 10	n.d.	oral, solution
	idiopathic, or due to noise	18		4 ml, 2x / day
	trauma, or to pressure	9 / 9		30 d.
	trauma			
	clinical + otoneurological exam, impedance, electronystagmography			
Meyer, 1986a	tinnitus:	259	53.2	Tanakan®
	ambulatory, <1 year;		137 / 122	usual dose and form
	not susceptible for	233		
	anti-infectious, surgical treatment	119 / 54 / 47		> 1 m.
Meyer, 1986b	tinnitus:	103	50.4	Tanakan®
	out-patients	58 / 45	54 / 49	oral, solution
	of recent origin: <1 year			4 ml/day
	not susceptible for anti-infectious, surgical treatment			1-13 m.
Haguenaer <i>et al.</i> , 1986	vertigo syndrome:	70	49.3	Tanakan®
	clinically confirmed	35 / 35	34 / 33	oral, solution
	recent onset (<2 yrs.)	67		160 mg/day (2x 2 ml)
	origin unknown	34 / 33		3 m.

Outcome measure	Result	Remarks
global effect (ENT specialist): '(very) good' / 'moderate' / 'failure'	7/1/1 vs 3/2/4	parallel vs nicergolin (NCG)
tonal audiometry (250, 500, 1000, 2000, 4000, 8000 Hz)	mean gain for 6 frequencies: +34.1 dB vs +23.0 dB +52.3% vs +37.8%	(alpha blocker) source: hospital (10 d.) + out-patient
vocal audiometry	Ginkgo: stronger effect	exclusion criteria
associated symptoms: tinnitus	7 vs 2 cured	
vestibular examination + caloric measurements	9 vs 3 normalized	
global assessment of development of tinnitus (physician): 6 categories (much improved — much deteriorated)	% cured or clearly improved: 63% vs 39% vs 30%	parallel vs nicergolin (NCG)
clinical assessment of tinnitus aspects: - intensity (scale 0-3) - discomfort (scale 0-3)	(prognostic factors evaluated) +0.98 vs +0.71 vs +0.54 (p<0.01) +0.94 vs +0.61 vs +0.51 (p<0.01)	+ vs almitrine-rau-basine (ARB) city practices of 15 ENT specialists
time needed to disappearance or clear reduction	75 vs 100 vs 156 days	pseudo-randomization (day of the week)
global assessment of change in clinical signs (timbre, periodicity, rhythm, discomfort); scale: 1-6 (1=much impaired, 2=impaired, 3=unchanged, 4=bit improved, 5=improved, 6=much improved, 7=treatment not tolerated)	% (much) improved: 56% vs 40% (p=0.05)	parallel vs placebo out-patients; city practices of 10 ENT-specialists
time until clear improvement / disappearance	70 vs 119 days (p=0.03)	exclusion criteria
change of intensity (scale 0-3)	+1.00 vs +0.67 (p=0.03)	prognostic factors:
change of level of discomfort (scale 0-3)	+0.84 vs +0.59 (p=0.08)	duration, location (uni- vs. bilateral), always vs intermittent
paraclinical criteria: tonal audiometry, impedance, electronystagmography		
global assessment of impact of vertigo on daily life; VAS (0-100 mm) (patient)	% improvement since onset: 75% vs 18% (p=0.01)	parallel vs placebo
global impression of improvement since start intervention; 5-point scale: disappeared, much improved, improved, stable, impaired (clinician)	% cured or clearly improved: 85% vs 42% (p<0.001)	3 centres exclusion criteria
questionnaire on intensity, frequency and duration of vertigo and associated signs (hearing loss, tinnitus, nausea, headache); 4-point scale: disappeared (0), diminished (1), stable (2), aggravated (3)	sum of scores w.1, w.2, w.3 (range: 0 - 9): intensity: +3.4 vs +4.5 (p=0.02) frequency: +3.2 vs +4.2 (p=0.02) duration: +3.3 vs +4.5 (p=0.03)	
ortoneurological examination: Romberg test, Babinski-Weill test (blind walking), nystagmus	% with normal values: +35% vs +16% (p=0.06)	



Author, year	Study population			Intervention
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Holgers <i>et al.</i> , 1994	permanent severe tinnitus (PST)	20	59.5 10 / 10	GBE (Seredrin®) oral, tablets 29.2 mg/day 2 w.
Hoffmann <i>et al.</i> , 1994	sudden hearing loss: - idiopathic - <11 days - no previous treatment	80 40 / 40	19-89 (40.5) 47 / 33	Tebonin® oral, solution + tablets 175 mg in 500 ml HAES 6% + 80 mg/day 3 w. HAES as standard therapy
Morgenstern and Biermann, 1997	tinnitus aurium: - chronic: 2 months; mean duration: 4.5 yrs. - reproducible	99 49 / 50	19-79 (45.5)	Tebonin® forte oral, tablets 120 mg/day 12 w.

Outcome measure	Result	Remarks
preference expressed: GBE vs. placebo subjective assessment of tinnitus: (VAS, 100 mm) <ul style="list-style-type: none"> - loudness - awareness - annoyance 	6 GBE, 7 placebo, 7 no pref. no. with strongest effect by: GBE/placebo/no sign. diff.: 2/1/15 (3 did not participate) 0/2/16 (3 did not participate) 2/2/14 (3 did not participate) inconsistency between drug preference and VAS recordings	cross-over vs placebo source: Hospital Audio- logy Dept., Göteborg RCT preceded by 2 w. open trial: 21 out of 80 patients who improved on GBE in open trial were admitted to RCT wash-out: 1 week
relative gain of hearing capacity	+23% vs +13% (p=0.06) total recovery, 1 w.: 40% vs 40% total recovery, 3 w.: 78% vs 55%	parallel vs naftidrofuryl (Dusodril®)
tinnitus: loudness (dB) (tonal audiometry)	+38 dB vs +34 dB	
tinnitus: intensity (subjective: VAS)	+4.3 vs +3.2 points	
hearing loss (sum of 250, 500 Hz, 1, 2, 4, 6, 8 KHz)	+180 dB vs +140 dB	
vestibular test		
<i>primary outcome measure:</i>		parallel
- tinnitus: loudness (most seriously affected ear) assessed by tonal audiometry	+3.3 vs -1.0 dB (p=0.015) (contralateral ear: no difference)	vs placebo source: ENT-clinic, ambulatory patients
<i>secondary outcome measures:</i>		run-in: 2 w. (placebo) drop-outs: 21 (13 vs 8) I-to-T analysis
- click-evoked otoacoustic emission (OAE)	no change	
- tinnitus: frequency and intensity		
- hearing loss (tonal audiometry)		
- subjective impression of tinnitus intensity by patient; scale 0-5	+0.4 vs +0.4	
- subjective impression of change in tinnitus complaints; scale: -5 - +5	% deteriorated / same / improved: 31%/61%/8% vs 14%/84%/2%	
- laboratory assessment (blood)		

Table 4. RCTs on the efficacy of *Ginkgo biloba* special extract in patients suffering from cerebral insufficiency: cognitive and non-cognitive symptoms.

Author, year	Study population		Intervention	
	Diagnostic category	Number randomized analyzed (EG/other)		Age (years) Sex (M/F)
Eckmann and Schlag, 1982	cerebral insufficiency: signs + symptoms neurological + psychical characteristics	50 25 / 25	45-74 30 / 20	Tebonin® forte oral, solution (drops) 120 mg (60 dr./day) 30 d.
Eckmann, 1990	cerebral insufficiency + depressive mood: ≥6 / 12 symptoms 'moderate' or 'severe'	60 58 29 / 29	41-71 31 / 29	Kaven® oral, solution 160 mg / day 6 w.
Brüchert <i>et al.</i> , 1991	cerebral insufficiency: typical symptoms HOPS: ≥1 symptom within 4 symptom groups present prevalence range of 11 symptoms: 45-95%	303 153 / 150 209 110 / 99	69 77 / 132	Kaven® (LI 1370) oral, tablets 150 mg/day 12 w.

Outcome measure	Result	Remarks
list of 11 'cerebrovascular' symptoms: 1. dizziness. 2. headache. 3. tinnitus. 4. awareness. 5. orientation. 6. motor function. 7. sensibility. 8. speaking: understanding. 9. speaking: ability. 10. depression. 11. anxiety	presence of each symptom: 1.+91 vs +43% 2.+92 vs +48% 3.+100 vs +50% 4.+92 vs +55% 5.+94 vs +53% 6.+92 vs +57% 7.+92 vs +55% 8.+93 vs +53% 9.+92 vs +43% 10.+92 vs +44% 11.+92 vs +44%	parallel vs placebo source: n.d. largest treatment effect: during days 8-18. median time until onset of improvement: 11 vs. >30 d. (p<0.01)
clinical assessment of neurological findings	+92% vs +44% (p<0.001)	
clinical assessment of psychopathology	+92% vs +44% (p<0.001)	
EEG		
laboratory tests (SGOT, SGPT, bilirubin, etc.)	no difference	
12 subjective clinical CI-symptoms: 1. dizziness. 2. ear noise. 3. headache. 4. confusion. 5. memory loss. 6. forgetfulness. 7. concentration. 8. fatigue. 9. loss of performance. 10. lack of motivation. 11. depression. 12. anxiety; 4 categories: absent, mild, moderate, severe	number of improved persons higher with Ginkgo vs. placebo, for all 12 symptoms (p.001); most improvements during weeks 2-4, e.g.: concentration: +23 vs +5 forgetfulness: +22 vs +5 depression: +24 vs +2 still improvement in weeks 4-6	parallel vs placebo hospitalized because of depression due to CI duration of symptoms: ca. 32 m.
global assessment of efficacy (patient) ('very good' / 'good' / 'moderate' / 'poor')	22/69/17/2 vs 8/43/34/14 (p<0.001)	parallel vs placebo
global assessment of efficacy (physician) ('good' / 'moderate' / 'unsatisfactory')	78/28/4 vs 31/29/39 (p<0.001)	out-patients 33 GP practices
11 subjective clinical CI-symptoms: memory, forgetfulness, concentration, fatigue, performance, motivation, depression, anxiety, dizziness, headache, tinnitus 4-point scale: severe, moderate, mild, absent	no. improved: Ginkgo vs. placebo: p <.01: 3 symptoms p <.05: 5 symptoms n.s., tendency: 1 symptom n.s.: 2 symptoms (anxiety, dizz.)	run-in period: 2 w. exclusion criteria
psychometric test: trail making test (figure connection test)	+20 vs +12 sec (p<0.01)	



Author, year	Study population		Intervention	
	Diagnostic category *	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Schmidt <i>et al.</i> , 1991	cerebral insufficiency: typical symptoms of reduced brain function typical symptoms HOPS: ≥1 symptom within each of 4 symptom groups present	100 99 50 / 49	51-68 45 / 54	Kaven® (LI 1370) n.d. 150 mg/day 12 w.
Hofferberth, 1991	organic brain psycho- syndrome (HOPS): stable condition	50 25 / 25	57-76 29 / 21	Kaven® (LI 1370) oral, tablets 150 mg/day 6 w.

Outcome measure	Result	Remarks
global assessment of efficacy (patient)	35/7/7 vs 7/12/30 (p<0.01)	parallel
global assessment of efficacy (physician) ('good' / 'moderate' / 'unsatisfactory')	36/6/8 vs 4/12/33 (p<0.01)	vs placebo out-patients:
12 subjective clinical CI-symptoms: memory, forgetfulness, concentration, fatigue, performance, motivation, depression, anxiety, dizziness, headache, tinnitus, confusion 4-point scale: severe, moderate, mild, absent; blood pressure, heart rate	no. improved: Ginkgo vs. placebo: p <0.01: 7 symptoms p <0.05: 1 symptom n.s., tendency: 2 symptoms n.s.: 2 symptoms (headache, confusion)	1 internist's practice mean duration of disease: 26 m. run-in period: 3 w. exclusion criteria
global effect (physician)	14 vs 0 good; 4 vs 11 moderate	parallel
11 clinical CI-symptoms (hetero-assessment): memory, forgetfulness, concentration, fatigue, performance, motivation, depression, anxiety, dizziness, headache, tinnitus; 4-point scale: severe, moderate, mild, absent	9 / 11 symptoms significant better with Ginkgo compared with placebo (p<0.01)	vs placebo hospital patients: neurological dept. of university hospital exclusion:
psychometric testing: - Vienna Determination Test (classic reaction test; 3 steps; abnormal result: <160, <168, <175 out of 180 correct) - number connection: >28 sec for 30 numbers	+18 vs -2 correct reactions (tempo step 10) (p<0.001) +12 sec vs +1 sec (p<0.001)	use of psychophar- maca, nootropics, vasoactive drugs run-in period: 2 w.
neurophysiological testing: - saccadic eye movements: duration of rapid eye movements (right) saccadic latency time (right) rapidity of eye movements (right) - EEG-analysis (theta part >70%) - evoked potential (P300 amplitude)	+55 msec vs +2 msec (p<0.001) +85 msec vs -5 msec (p<0.001) +150 vs +15 ° / s. (p<0.001) +20% vs +5% (p<0.001) +7.5 vs +0.0 microV (p<0.001)	

Table 5. RCTs on the efficacy of *Ginkgo biloba* special extract in patients suffering from cerebral insufficiency: depressive mood as a major sign.

Author, year	Study population			Intervention
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Halama, 1990	cerebral insufficiency + depression: - mild to moderate CI - non-responder to anti-depressants (>3 m.)	20	55-85 9 / 11	Tebonin® forte oral, tablets 240 mg/day 8 w.
Stocksmeier and Eberlein, 1992	cerebral insufficiency + depressive mood (mild): - subjective experience + clinical diagnosis - HAMD-score >15 - KLT-score <90% of norm	60 30 / 30 36 17 / 19	51-78 17 / 43	Tebonin® forte oral, tablets 120 mg/day 12 w.
Schubert and Halama, 1993	cerebral dysfunction (mild, moderate) + depression: - no clear effect of anti-depressants (3 m.) - CI-scale >20 points	40 20 / 20	50-78 13 / 27	Tebonin® forte oral, tablets 240 mg/day 8 w.

Outcome measure	Result	Remarks
<i>primary outcome measures:</i>		parallel
- Zung's Self-Rating Depression Scale (SDS)	+5.3 vs -5.2	vs placebo
- late acoustic-evoked potential (SAEP/ALEP):		anti-depressive
latency-time P3-wave	+13.3 vs -9.8 msec (p=0.048)	regimen continued
latency-time CNV-wave	+46.4 vs -26.8 msec	pilot study
<i>secondary outcome measures:</i>		
- CI-test (Lehrl et al): interference	+11.7 vs -0.1	
- CI-test (Lehrl et al): symbols, numbers, etc.	+3.4 vs -1.0	
- CI-scale (Weidenhammer & Fischer)	+3.1 vs -1.9	
- Kurztest für allgemeine Intelligenz (KAI)	+2.5 vs -0.6	
- subjective assessment of depression	3 vs 0 improved	
- subjective assessment of concentration	verum group improved	
<i>primary outcome measure:</i>		parallel
- Hamilton Depression Scale (HAMD), modified, sum score items 1-17	+6.0 vs +4.6 (baseline: 13.3 vs 15.0)	vs placebo source: n.d. MID, AD excluded
<i>secondary outcome measures:</i>		
- Konzentrations-Leistungs-Test (KLT)	n.s.	
- Gießener well-being questionnaire	n.s.	
- newly developed computerized test	n.s.	
<i>primary outcome measure:</i>		parallel
- Hamilton Depression Scale (HAMD), modified, sum score items 1-17	4.5 vs 13.0 (p < 0.01) (baseline: 14.0 vs 14.0) >+5.0: 15 / 20 vs 6 / 20	EGb 761 + anti-depressants vs placebo
<i>secondary outcome measures:</i>		out-patients neuro- psychiatric practice
- CI-test (Lehrl & Fischer)	n.s.	
- CI-scale (Weidenhammer & Fischer)	n.s.	excluded: serious
- Kurztest für allgemeine Intelligenz (KAI)	+15 vs +4 points (p=0.02)	dementia, endoge- neous depression, etc.
- severity of symptoms (assessed by physician):	no. moderate - severe:	
depression (5 categories)	19 → 4 vs 20 → 15	
concentration (5 categories)	19 → 8 vs 20 → 12	
- late acoustic evoked potential (SAEP)	n.s.	

Table 6. RCTs on the efficacy of *Ginkgo biloba* special extract in patients suffering from cerebral insufficiency: unspecified.

Author, year	Study population	Intervention
	Diagnostic category	Brand name
	Number randomized analyzed (EG/other)	Administration route
		Dose
		Duration
Moreau, 1975	cerebral insufficiency	Tanakan® (Ipsen 225)
	diagnostic criteria: n.d.	n.d.
		120 mg/day
		3 m.
Bono and Mouren, 1975	cerebral insufficiency:	Tanakan®
	manifestations: attention, mood, memory, confusion	oral, solution (drink)
	functional signs: vertigo, visus, headache, tinnitus, motoric + sensoric, etc.	120 mg/day (3 ml)
		5-6 w.
Augustin, 1976	hospital patients with atherosclerosis, and often signs of vascular origin (cardiac, cerebral, ocular, renal, etc.)	Tanakan®
		n.d.
		120 mg/day
		6 m.

Outcome measure	Result	Remarks
global effect (% excellent or good)	79% vs 21% ($p < 0.001$)	parallel
physical signs + functional symptoms:	% improved:	vs placebo
- vertigo / walking instability	91% vs 50%	source: n.d.
- walking activity	75% vs 42%	also comparison with
- sensoric manifestations	75% vs 13%	ergot alkaloids (n=30)
- headache	78% vs 38%	
- temporo-spatial desorientation	94% vs 31%	
- walking distance	67% vs 17%	
general activities (impact of disease on daily life activities (sleeping, appetite, mood, etc.))	% improved:	
psychometric tests: 10 tests (temporo-spatial orientation, memory, etc.)	90% vs 43%	
	significantly more improvement ($p < 0.05$) in Ginkgo group for 6 / 10 tests	
global effect (good / moderate / no)	8/3/3 vs 3/2/9 ($p < 0.01$)	parallel
evolution of functional signs (vertigo, tinnitus, headache, memory, visuo-spatial orientation, confusion, visual / auditive / walking problems)	vertigo, walking problems, headache, visual problems, tinnitus: 90% disappeared	vs placebo
evolution of neurological signs (motor or sensoric defects, language problems)	memory, confusion, orientation: 62% disappeared	source: n.d.
evolution of behavioural signs (vigilance, sociability, mood, sleep)	motor defects: 65% improved	also GBSE - 3 ergot
clinical examination (cardiovascular, renal, etc.)	sensoric defects: 40% improved	derivates comparison
paraclinical examination (EEG, ECG, eyes, blood)	vigilance: >70% recovered	(20 vs 20)
global assessment of activities (reading, interest in radio / TV, usual manual activities)	improved: 44% vs 13% ($p = 0.001$)	parallel
psychometric test battery:	n = 168; sum score of 5 tests:	vs placebo
1. orientation. 2. intellectual efficiency.	- no. improved: 61 vs 26	hospital patients
3. short term memory. 4. Rey test (learning words). 5. Benton visual retention test	- no. aggravated: 13 vs 36	(months, years)
clinical assessment of vascular state:		same nutrition, dependent
- arterial blood pressure	n.s.	prestratification: age
- oscillometry: arterial peripheral circulation	index increased: 24 vs 2	
- central retinal artery pressure	n.s.	
- atheromatous retinopathy	3 improved vs 6 impaired	
biological surveillance, morbidity/mortality	0 vs 8 deaths	



Author, year	Study population			Intervention
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Israëli <i>et al.</i> , 1977	intellectual deficiency syndrome caused by insufficient cerebral circulation; diagnostic criteria: n.d.	60 30 / 30 48 24 / 24	62-96 (72) 32 / 16	Tanakan® oral, solution 240 mg/day (6 ml) 60 d.
Agnoli, 1980	cerebral insufficiency: stable symptoms, >30 d. PBRSM-score ≥30 Plutchik GRS ≥16	30	51-83 (65.3) 16 / 14	Tebonin® forte oral, solution 150 drops/day 4 w.
Dieli <i>et al.</i> , 1981	cerebral insufficiency: chronic signs, symptoms HIS-score <7 Plutchik GRS	40 20 / 20 25 11 / 14	62	Tanakan® oral, solution 3 ml/day 60 d.
Haan <i>et al.</i> , 1982	cerebral insufficiency: signs + symptoms: >11	60 30 / 30	45-65 (54.0) 60 / 0	Tebonin® forte i.v. (infusion) 97.5 mg/day (2 h.) 15 d.
Pidoux <i>et al.</i> , 1983	senile deterioration: EEG-abnormalities (slow dominant frequency at 8 - 8.5 cps, no focalized abnormalities / asymmetry)	14 7 / 7 12 7 / 5	84 0 / 12	GBE n.d. 160 mg/day 12 w.

Outcome measure	Result	Remarks
Crichton scale (modified, adapted)		parallel
4 factors derived from 7 psychometric tests (code test, verbal fluency, numbers repeat, etc.):		vs placebo
1. attention (vigilance) + concentration	+0.44 vs +0.21	source: n.d.
2. memory	+0.38 vs +0.13	
3. fluency of operant behaviour	+0.00 vs -0.08	
4. effective activity + psychomotor efficiency	+0.04 vs +0.13	
global clinical assessment	23x Ginkgo vs 3x placebo better	cross-over
Parkside Behaviour Rating Scale, modified (PBRSM):16 items (1-5), 8 subcategories: vigilance, orientation, memory, language behaviour, adjustment, socialization, autonomy	Ginkgo vs placebo: +85% vs +1% change change (borderline) significant: 7 vs 0 items (not: anxiety)	vs placebo source: n.d.
Plutchik Geriatric Rating Scale (GRS): 31 items (Q, 1, 2)	+84% vs +1% change	
global clinical assessment (physician): excellent / good / moderate / none	8/7/1/4 vs 3/4/0/13 (p.05)	parallel vs placebo
Plutchik Geriatric Rating Scale (31 items)	13 vs 3 items sign. improvement; Ginkgo sign. better on 11 items	hospital patients exclusion criteria
memory of numbers (Wechsler Bellevue)	+4.46 vs +0.57 (p=0.01)	
code test (Wechsler Bellevue)	+17.73 vs +4.29 (p=0.02)	
clinical examination		
laboratory examinations		
psychological tests:		parallel
- concentration: d2-test Brickenkamp	394 vs 374 items (tendency)	vs vincamine
- short term memory: digit span (HAWIE)	71 / 321 vs 69 / 322 sec	(vasoactive agent)
- visual-motor test: mosaic test (WIP)	7.2 / 5.3 vs 6.7 / 5.0 numbers	source: n.d.
cranial computer tomogram (CCT)	only as part of the screening	exclusion criteria
neurological examination	3/10 vs 4/10 improved	
EKG + internal medical examination	no changes	
EEG	18 (69%) vs 12 (48%) improved	
laboratory tests	no changes	
Doppler, angiography aorta (if indicated)		
clinical rating scales:	+29% vs +13%	parallel
- SCAG		vs placebo
- SGRS		source: n.d.
clinical examination		
EEG patterns	improvement GBE: increase alpha rhythm, reduction theta	

Author, year	Study population			Intervention
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Teigeler and Pieprzyk, 1984	cerebral insufficiency: slight to moderate screening procedure premorbid IQ >80 (MWT-B)	40 20 / 20	n.d. n.d.	Tebonin® forte oral, tablets n.d. (4 tablets / day) 4 w.
Geßner <i>et al.</i> , 1985	cerebral deterioration: symptoms reported by patients	40 20 / 20	57-77 (65.5)	rökan®, Tanakan® n.d. 120 mg/day 12 w.
Arrigo, 1986 Arrigo and Cattaneo, 1985	cerebrovascular insufficiency: chronic vertigo, headache, tinnitus, loss of memory, learning ability	90 80	36-88 (66.5) 59 / 31	Tebonin® forte oral, solution (drops) 120 mg/day 45 d.
Taillandier <i>et al.</i> , 1986	cerebral disorders due to ageing: EACG score: - 3-5 for at least 2 / 6 items - not 7 for >2 / 6 items - total score: 21-113 (moderate cognitive deficiency) HIS-score: 0-11	166 80 / 86	60-97 (82.1) 26 / 140	Tanakan® oral, solution 160 mg/day (2x 2 ml) 12 m.

Outcome measure	Result	Remarks
information processing properties:		parallel
1. short-term storage capacity (mental efficiency)	+24.8 vs +9.8 bit/sec +45% vs +13%	vs placebo
2. learning speed (learning efficiency)	evaluated only for placebo	hospital admission: orthopedic + internal treatment: 2 w. in hos- pital + 2 w. at home
general efficacy (patient): % (very) good	50% vs 27%	parallel
general efficacy (investigator): % (very) good	44% vs 19%	vs placebo
EEG-analysis: resting EEG + vigilance EEG	only increase in vigilance on resting EEG, if low initial theta/ alpha-ratio and dominant freq.	volunteers, relatively healthy also comparison with nicergolin (n=20)
psycho(physio)logical measures:		wash-out: 2 w.
- simple reaction time (light signal)	+20 msec vs +0 msec (n.s.)	
- multiple choice reaction test (MCRT; Vienna Determination Apparatus)	+8 msec vs -5 msec (n.s.)	
- flicker fusion frequency measurement	n.s.	
- pursuit tracking test	n.s.	
- Pauli test (simple calculations)	n.s.; GBE faster improvement	
- d2 test	n.s.	
physiological measures: EEG, BP, pulse rate	n.s.	
clinical symptoms (self-rating questionnaire): 26 questions (+, 0, -), 7 symptoms: 1. asthenia 2. headache 3. tinnitus 4. dizziness 5. insomnia 6. visual disturbances 7. tension	change in number (groups A+B): 1.+22 vs -6 2.+16 vs +2. 3.+18 vs +10 4.+38 vs +8 5.+10 vs -2 6.+12 vs +4 7.+26 vs +18.	cross-over vs placebo source: n.d. prestratification (4 categories: WAIS)
WAIS: psychometric testing: 4 tests not sensitive to mental loss: 1, 6, 8, 11 4 tests sensitive to mental loss: 4, 5, 7, 9	EGB + plac. vs. plac. + EGB: +0.3 vs +0.5 (n.s.) +0.2 vs +0.8 (p<0.04)	run-in period: 15 d. wash-out: 15 d.
cube test (WAIS): concentration, orientation	+0.7 vs +0.2 (p=0.017) (period 1)	
word recognition	+16.3 vs +2.6 (p<0.001) (period 1)	
Rey Figure test (shape B)	+2.4 vs +1.4 (p<0.05) (period 1)	
memory performance (reproducing objects)	+15.5 vs +5.7 (p<0.001) (period 1)	
anxiety (state; STAI questionnaire)	+5.6 vs +2.5 (p<0.001) (period 1)	
anxious behaviour (trait; STAI questionnaire)	+2.8 vs +1.4 (p<0.01) (period 1)	
geriatric clinical evaluation scale (EACG): 17 items: 1-7 (absent - very important)	+9.8 vs +5.2 (p=0.01)	parallel vs placebo
general clinical assessment of state: 6-point scale: absent, mild, moderate, average, important, serious (clinician)	% 'mild': +12% vs -7%	4 homes for the elderly (>2 m. inhabitant)
global assessment of evolution; 6-point scale: much improved, improved, little improved, unchanged, bit impaired, impaired	58% vs 43% improved	



Author, year	Study population			Intervention
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Halama <i>et al.</i> , 1988	cerebrovascular insufficiency (mild to moderate) clinical findings: HIS >7 psychometric tests + rating scales: Crichton Geriatric Scale: stages 1-3	40 20 / 20	55-85 (65.9) 23 / 17	Tebonin® forte oral, tablets 120 mg/day 12 w.
Hofferberth, 1989	cerebro-organic psychosyndrome: - clinical symptoms (e.g. dizziness, memory) - 4 screening tests (EEG, saccadic eye movements, Wiener Determination test, number connection) ≥2 abnormal results	36 18 / 18	53-69 (63.2) 23 / 13	rokan® oral, tablets 120 mg/day 8 w.
Vorberg <i>et al.</i> , 1989	cerebral insufficiency: presence of clinical symptoms (≥4 out of 6)	100 96 49 / 47	55-85 (70) 60 / 40	Kaveri® solution oral, solution 112 mg/day 12 w. 27-30% flavonoids + 1-2% terpenetriactones
Gerhardt <i>et al.</i> , 1990	cerebrovascular disorders (HOPS, MID, SDAT) symptoms: dizziness (31 vs 31), concentration (6 vs 7), memory (8 vs 7), headache (8 vs 5), depression (7 vs 4) Hachinsky Ischemic Score	80 40 / 40 65 32 / 33	55-88 (71) 23 / 42	EGb 761 oral, solution (3 ml) 120 mg/day 6 w.

Outcome measure	Result	Remarks
<i>primary outcome measure:</i>		parallel
- SCAG: 5 factors; 18 items; item score: 1-7; total score: 18-126	+9 vs +0 points ($p < 0.0001$) esp.: STM, dizziness; etc.	vs. placebo out-patients (neuro- logical practice)
<i>secondary outcome measures:</i>		exclusion criteria: a.o.: Alzheimer
- Crichton scale (CRBRS)	n.s.	
- Syndrome Kurz Test (SKT)	+3.5 vs +0.8	
- symptom rating: 1. dizziness. 2. ear noise. 3. headache. 4. hearing loss; 4 categories	1.+1.4 vs +0.1 2.+0.6 vs +0.15 3.+0.9 vs +0.1 4. n.s.	
- background interference measurement	n.s.	
- craniocorponography (CCG)	n.s.	
overall effect: 4 categories (patient)	9 / 6 vs 2 / 0 much impr./improved	parallel
overall effect: 4 categories (physician)	6 / 8 vs 0 / 2 much impr./improved	vs placebo
psychometric tests:		hospitalized patients
- Wiener Determination Test (3 steps; <160, <168, <175 out of 180 possible correct eractions)	step 10: +19 vs -1 ($p < 0.0001$) step 12: +14 vs -2 ($p < 0.0001$) step 15: +6 vs +2 ($p < 0.0001$)	block randomization (6 patients)
- number connection test (28 sec. needed to connect 30 numbers)	+13 vs +0 sec ($p < 0.0001$)	wash-out: 2 weeks exclusion criteria
neurophysiological tests:		
- saccadic eye movements: duration (abnormal: >150 msec) latency (abnormal: >250 msec) rapidity (abnormal: <250 %/sec)	+45 vs -2 msec ($p < 0.0001$) +53 vs -8 msec ($p < 0.0001$)	
- quantified EEG (theta-part >70%)	+13% vs -1% ($p < 0.0001$)	
list of 6 clinical symptoms:		parallel
1. memory. 2. concentration. 3. anxiety.	1. +1.4 vs +0.5 ($p < 0.001$)	vs placebo
4. dizziness. 5. headache. 6. ear noise.	2. +1.4 vs +0.4 ($p < 0.001$) 3. +1.1 vs +0.4 ($p < 0.001$)	ambulatory patients (GP-practice)
degrees of severity: 1-5	4. +2.0 vs +0.6 ($p < 0.001$)	exclusion criteria
in accordance with modified Crichton scale	5. +2.3 vs +0.9 ($p < 0.001$)	duration of complaints: ca. 20 m.
pulse rate, blood pressure	6. +0.8 vs +0.4 (n.s., tendency)	
Nuremberg Elderly Rating (NAR):		parallel
hetero-assessment personality profile, 9 pairs of items (item score: 1 - 7)	1 vs 3 items sign. improved Ginkgo: careful - not careful	vs dihydroergotoxin ambulatory patients
personality profile: self-assessment, 7 pairs of items (scale: VAS, 0-22 cm)	5 vs 7 items improved; Ginkgo: 2 impaired; changes n.s.	single-blind (patient) MID, DAT not dis- tinguished (HIS)
psychomotoric coordination:		
- trail making test	total time: +3 vs +10 sec (n.s.)	
- digit symbol test	+2 vs +4 points (n.s.)	



Author, year	Study population			Intervention
	Diagnostic category	Number randomized analyzed (EG/other)	Age (years) Sex (M/F)	Brand name Administration route Dose Duration
Maier-Hauff, 1991	cerebral insufficiency after subarachnoid haemorrhage (SAH) + surgery	50 24 / 26 42 20 / 22	21-63 (48) 26 / 24	Kaven® (LI 1370) oral, tablets 150 mg/day 12 w.
Gräbel, 1992	cerebral insufficiency: 1. patient history; indications for intracranial stenosis process (TIA, hypertension, etc.) 2. Doppler sonography (in case of doubt) 3. IQ-STM < IQ-MWT	72 36 / 36 53 29 / 24	63.8 35 / 37	EGb 761 oral, solid form 160 mg/day 24 w.

3

Efficacy of *Ginkgo biloba* in elderly people with dementia or age-associated memory impairment: design of a randomized clinical trial

Martien CJM van Dongen, MSc¹

Erik van Rossum, PhD¹,

Paul G Knipschild, MD, PhD²

Hilde Sielhorst, RN^{1,2}

Alphons GH Kessels, MD¹

¹Department of Epidemiology, Maastricht University

²Department of General Practice, Maastricht University

ABSTRACT

Since 1975 many controlled trials have investigated the efficacy and safety of *Ginkgo biloba* special extract in patients with various manifestations of cerebral insufficiency. Most of these trials showed very promising results, which point to a beneficial effect of Ginkgo treatment. The majority of the trials can be criticized, however, because of methodological flaws and fallacies. In order to substantiate the available evidence new research efforts are needed. The Maastricht Ginkgo Trial was initiated to help settle the dispute about the efficacy of Ginkgo in patients with dementia and cognitive decline, and to answer additional questions with regard to the optimal treatment dose and the persistency of any treatment effect. This paper presents the outline of the Maastricht Ginkgo Trial, as well as the rationale of the choices that make up the design of this trial. Successively, it deals with the choice of the study population (source population, recruitment strategy, eligibility criteria, sample size), the study treatment, and the study outcome measures, and the design measures taken to strengthen the comparability of the distinct study subgroups, the intervention conditions, and the outcome assessment procedures. Moreover, it addresses the procedures for data management and analysis.

1 INTRODUCTION

In the early 1990s the efficacy of *Ginkgo biloba* special extracts had been evaluated already in an impressive number of intervention trials. In 1992 a systematic review was published, that summarized the results and the methodological characteristics of 40 randomized clinical trials with Ginkgo in patients suffering from cerebral insufficiency¹. The review revealed that almost all retrieved trial reports showed a beneficial effect of Ginkgo. The methodological quality of the studies included in the review appeared to diverge considerably, with total scores varying from 18 to 90 on a 100-point assessment scale. An inadequate sample size, an insufficient description of randomization procedures, patient characteristics and outcome measures, questionable blinding procedures, and an incomplete presentation of the study results were identified as frequently occurring shortcomings.

Owing to the unsatisfactory methodological quality of many previous Ginkgo trials, conclusions about the efficacy of Ginkgo in cerebral insufficiency should be drawn with caution, and, therefore, the subject remains controversial. We initiated the Maastricht Ginkgo Trial in 1993, in order to provide new evidence for the efficacy of Ginkgo, based on valid and precise observations. Anticipating a positive effect of Ginkgo, we decided to design this trial in such a way that it would also shed light on the dose-effect relationship and the persistency of any Ginkgo effect. Moreover, we decided that the trial should focus on cerebral insufficiency patients who fulfilled rather well-accepted diagnostic criteria for cognitive decline: patients with dementia and non-demented patients with age-associated memory impairment (AAMI).

The Maastricht Ginkgo Trial was designed and planned during 1993. It was conducted (intervention, data collection) during the period 1994-1996. The analysis of the study

data took place in 1997. The trial was commissioned and financially supported by Dr. Willmar Schwabe Arzneimittel GmbH & Co, Karlsruhe, Germany (study number: 523001.01.016).

The current description of the design of the Maastricht Ginkgo Trial is largely based on the official study protocol. This protocol was approved and ratified by the members of the project team (Department of Epidemiology, Maastricht University, the Netherlands) and legal representatives of the Schwabe company before the start of the trial, in January 1994. Nevertheless we will describe the trial profile in retrospect. The same protocol was also part of the documentation that was assessed by the combined Medical-Ethical Committee of the Maastricht University Hospital and the Maastricht University, Faculties of Medicine and Health Sciences. In december 1993 this authority declared to have no objections against the trial. Some minor amendments to the initial protocol were made after the start of the trial, in November 1994 - concerning the admission criteria for dementia and AAMI patients (range of eligible SKT baseline scores) - and in July 1996 - concerning a change in the treatment allocation ratio and a specification of the hypothesis testing procedures within the scope of the confirmatory analysis of the trial results. We have incorporated these amendments, which have also been ratified officially, in the current design description. Furthermore, we have added some information that relates to the actual realization of the trial plan, in particular with regard to patient recruitment (numbers enrolled) and distribution (numbers allocated to each of the trial arms).

2 GENERAL AIM AND OUTLINE OF THE CURRENT STUDY

The aim of the Maastricht Ginkgo Trial was to assess the effects of varying amounts of

therapeutic dosage and varying duration of intervention with orally administered *Ginkgo biloba* special extract (EGb 761) on the occurrence of symptoms that indicate memory impairment or dementia (mild to moderate) in elderly people. The study was designed and carried out in accordance with the guidelines for good clinical practice (GCP) that have been developed for trials on medicinal products in the European community²⁻⁴.

In view of the design, the study can be characterized as a randomized, double-blind, placebo-controlled, parallel-group trial. After a 3-week qualification period (run-in phase) those subjects from the source population who were considered eligible for the study population, were randomly allocated to three parallel treatment regimens. One subgroup received the 'usual' dosage of EGb 761 (160 mg daily; 'verum group 1: usual dose'), a second subgroup was treated with an elevated dosage of EGb 761 (240 mg daily; 'verum group 2: elevated dose'), and the remaining part of the study participants was exposed to a placebo preparation, containing no EGb 761 at all ('placebo group').

The initial study medication was sustained for 12 weeks (first intervention period). The study subjects who completed this 12-week period were subsequently treated for another 12 weeks (second intervention period). At the beginning of this second intervention phase the 'verum group 1 (usual dose)' was split up randomly once more: half of its members moved to placebo treatment until the end of the follow-up period, whereas the other half continued on usual-dose EGb 761. The 'verum group 2: elevated dose' was subdivided in a similar way: half of the participants shifted to placebo treatment, and half of them continued with EGb 761. The original regimen of the 'placebo group' was maintained during the rest of the study. At the end of the second intervention phase - after 24 weeks of intervention - the efficacy of EGb 761 was assessed. The second intervention

phase also served the purpose of evaluating the durability of the intervention effect.

Data on outcome measures were collected at the end of the first intervention period - 12 weeks after the start of the trial - and at the end of the second intervention period - 24 weeks after the start of the trial - respectively. In addition, some of the outcome phenomena were also assessed in between, at regular intervals (after 4, 8, and 18 weeks of intervention). Figure 1 displays the outline of the study design.

In order to cover all relevant details of the study design we will successively address the following topics: study population (section 3), study treatment (section 4), study outcomes (section 5), and data management and analysis (section 6).

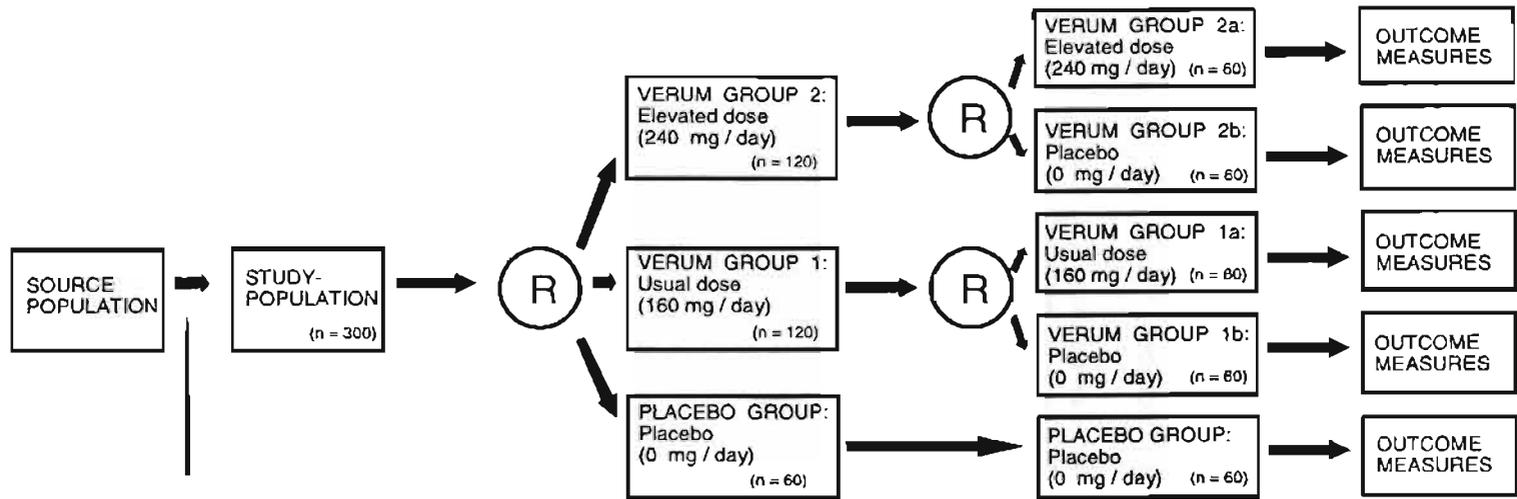
3 STUDY POPULATION

3.1 CHOICE OF STUDY PARTICIPANTS

3.1.1 Source population and recruitment of study candidates

The patient recruitment procedure aimed at the identification and selection of elderly people suffering from the dementia syndrome in a mild to moderate degree as well as elderly people with non-dementia stages of cognitive function impairment.

We decided to depart from residents of homes for the elderly in the Netherlands as the source population for the recruitment of eligible study subjects. These old people's homes provide sheltered living conditions for elderly persons who are not able anymore to live independently. Each resident occupies a (small) apartment in the home. Cleaning and the preparation of the meals is taken care of by the house staff. If necessary, residents receive personal or nursing care from the house staff, while the general practitioner re-



Eligibility assessment

- Inclusion criteria
- Exclusion criteria
- Informed consent

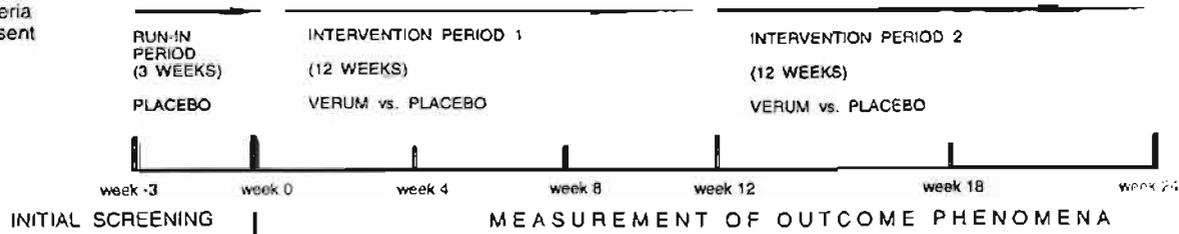


Figure 1. Design of the Maastricht Ginkgo Trial.

mains responsible for medical care. In the beginning of the 1990s the average age of residents of old people's homes in the Netherlands was approximately 84 years at admission.

We observed several advantages of these old people's homes as a source for the recruitment of anti-dementia drug trial participants. Firstly, these institutions are supposed to accommodate a relatively high proportion of eligible candidates at a central place. Secondly, the administration of the study medication could be incorporated within the daily routine of drug supply. The same applies to the monitoring of the compliance with the prescribed drug regimen. Thirdly, additional information could easily be made available, for instance with regard to the pattern of previous and current medicine use, medical history, results of clinical observations by caregivers (complaints, behaviour, activities of daily life, etc.). The choice of homes for the elderly as the sampling frame for our trial was also motivated by logistic considerations (easy to distribute study medication and to arrange appointments for screening and follow-up interviews), and by the wish to avoid an overrepresentation of 'hypochondriacs', viz. people who needlessly worry about cognitive decline, and who may disproportionately be attracted by a recruitment strategy based on advertisements aimed at the open population.

Initially, the selection of participants was restricted to the southern part of the province of Limburg in the Netherlands. When the trial was in progress the recruitment area was extended to the northern part of Limburg as well as to a region in the eastern part of the province of Noord-Brabant. A total of 39 homes participated in the study. In 7 of these homes the recruitment procedure was performed twice, with ample time in between. We expected that a considerable amount of

new cases would have occurred in the meantime. The number of residents ranged from about 50 to 200 residents per home, with an average number of approximately 100. Out of a total population of 4,459 residents in the 39 homes, 214 passed the selection procedure. They were randomly allocated to the various treatment alternatives. Details of the selection procedure are described in section 3.1.4 of this paper.

3.1.2 Eligibility criteria

Inclusion criteria

The following inclusion criteria were applied to guide the selection of the study participants:

1. Mental state: either diagnosis of (mild or moderate) dementia, according to DSM-III-R criteria, or presence of age-associated memory impairment (AAMI), according to adapted AAMI criteria, which refer to objectively measured memory decline in combination with subjectively perceived memory complaints. These criteria regarding the mental state will be elaborated below.
2. Written informed consent and agreement to comply with the trial conditions (informed consent declaration signed by the study subject or by a legal deputy). In each case the attending general practitioner was asked to check whether on medical grounds any objections existed against admission to the trial.

Mild to moderate dementia

In order to be classified as a dementia case (mild / moderate degree of severity) study applicants had to meet the DSM-III-R criteria for the diagnosis 'dementia'. The SIDAM ('Strukturiertes Interview für die Diagnose einer Demenz vom Alzheimer Typ, Multiinfarktdemenz und Demenzen anderer Ätiologie') was applied for the detection and differential diagnosis of dementia. The SI-

DAM has been conceived as a quick screening tool to determine whether dementia is likely to be present and, if yes, which type of dementia and which degree of severity it concerns. This happens in accordance with both the DSM-III-R and the ICD-10 criteria for dementia^{5,6}. The SIDAM can be seen as a translation of these criteria into operational terms^{7,8}.

The SIDAM consists of two parts. The first part - the test part - consists of questions and tests to be presented to the patients, whereas the second part involves clinical assessment. The first part is related to the DSM-III-R criteria A through B3, which refer to the presence of demonstrable evidence of impairment of the short-term and the long-term memory, impairment of abstract thinking capacity, impairment of judgment abilities, and any other disturbances of higher cortical function (e.g., language, motor skills, perception, ability to perform construction tasks). The second part of the SIDAM is connected with the DSM-III-R dementia criteria B4 through E, which refer to the presence of a change in personality traits, interference of the cognitive decline with work-related and social activities, the manifestation of cognitive dysfunction not only in a state of impaired consciousness, and the existence of indications for the presence of a specific organic cause or the absence of a non-organic cause, e.g., depression, that underlies the cognitive problems. The measurements that make up part one of the SIDAM were conducted by members of the research team. The questions that relate to criterion A (short- and long-term memory), include the Mini-Mental State Examination (MMSE)⁹. For this purpose the standardized version of the MMSE was used. The clinical assessment part of the SIDAM (part two) was administered by staff members of the participating homes for the elderly. One of the elements of this part was a series of questions yielding the so-called Hachinski

Ischemic Score (HIS), a measure aimed at the differential diagnosis of Alzheimer-type dementia (AD) and vascular dementia (VD)^{10,11}. As not all original HIS-items seemed to be equally relevant and easy to interpret, we decided to use a slightly modified version in the current trial. Basic information regarding the medical history and previous diagnoses of the study participants could be retrieved from the archives in the participating homes. If necessary, the responsible general practitioner was contacted to provide additional information.

Based on the SIDAM-results only persons with a mild or moderate form of Alzheimer-type dementia (AD), vascular dementia (VD; previously called multi-infarction dementia (MID)), or a mixed form of both types, were included in the trial. People suffering from dementia associated with another somatic cause or an unspecified type of dementia, as well as people with a severe state of dementia were excluded. In addition to the SIDAM findings, the results of the Syndrom Kurz Test (SKT) - a short test of memory and attention, see below - were used to demarcate the screened population according to the degree of severity of the cognitive dysfunction (range of eligible scores for demented patients: initially 9 through 23, later 8 through 23).

Age-associated memory impairment (AAMI)

To be identified as a case of age-associated memory impairment (AAMI) study candidates had to meet the following criteria:

1. Presence of an impaired level of cognitive performance, as indicated by a test score from 8 through 23 (initially: 9 through 18) on the Syndrom Kurz Test (SKT).
2. Presence of serious self-reported complaints of memory loss in everyday life, as indicated by a score ≥ 12 on the Memory Assessment Clinics Questionnaire (MAC-Q).

3. Presence of an adequate level of pre-morbid intellectual capacity.
4. Absence of dementia.

The SKT is a brief cognitive performance test for the assessment of memory and attention. It mainly tests perceptual-motor speed, memory and language (also with a speed component), not orientation and basic performance abilities. The SKT consists of 9 subtests. Six of them are speed-oriented: naming objects (assessment of language fluency), naming numerals (assessment of number fluency) arranging blocks with two-digit numerals (assessment of attention, planning, and praxis), replacing blocks (assessment of short-term memory, attention, planning and praxis), counting symbols (assessment of attention and concentration), and reversal naming (assessment of attention and concentration, in particular susceptibility for interference). The three remaining tests are span-oriented: immediate recall of objects (assessment of short-term visual memory), delayed recall of objects (assessment of long-term memory span), and recognition of objects (assessment of recognition visual memory span). It takes 10-15 minutes to complete all subtests¹².

The SKT has been extensively validated. Its factor structure (two factors: 'speed of reaction' or 'attention', and 'memory'), reliability (test-retest correlation), content validity, cross-sectional or concurrent validity (correlations of SKT-scores with scores on other psychometric test batteries, global and detailed clinical assessment (e.g., Clinical Global Impression (CGI), MMSE, Alzheimer's Disease Assessment Scale (ADAS)), and laboratory examinations (e.g., results of computer tomography (CT)), and longitudinal validity ('sensitivity to change') have been evaluated thoroughly, with satisfying findings¹³⁻¹⁷. Reference scores are available for combined categories of age (4x) and level of premorbid intelligence (3x).

Based on these reference scores the raw scores for each subtest (seconds or numbers) can be easily converted into scaled scores, resulting in a total SKT-score varying from 0 (most beneficial) to 27. In our trial we derived scaled SKT-scores, assuming an 'average level of intelligence' for each study participant (IQ-score between 90 and 110). It is claimed that the SKT is able to differentiate within the spectrum of mild to severe cognition disorders. The total SKT-score is supposed to correspond with the prevailing clinical picture of mental deterioration as follows:

- score 0-4: no indications for the presence of cerebral function insufficiency in the sense of a (brain-)organic psychosyndrome or an incipient demential disorder
- score 5-8: indications for the presence of a (brain-)organic psychosyndrome or a beginning demential disorder
- score 9-13: mild (brain-)organic psychosyndrome; starting demential symptoms accompanied with mild cognitive function disturbances
- score 14-18: moderate (brain-)organic psychosyndrome; moderate expression of cognitive function deficiency and symptoms of dementia
- score 19-23: severe (brain-)organic psychosyndrome; pronounced cognitive function deficiencies and symptoms of dementia
- score 24-27: severe cognitive function defects, typical for the clinical picture of an advanced stage of dementia.

The Memory Assessment Clinics Questionnaire (MAC-Q) was used in our trial to assess self-perceived memory functioning^{18,19}. For this purpose the original MAC-Q was slightly modified and adapted to match better with the characteristics and living conditions of the target population. Five questions were asked, which covered the following topics: remembering names, orientation

in a familiar environment, forgetting appointments and activities, losing track of personal belongings, and forgetting things one just has read or seen. Each question was rated at a 5-point response scale (extreme scores: 1 = "never a problem"; 5 = "always a problem"). Based on the item-scores a total score was calculated, which ranged from 5 to 25 points. The MAC-Q was rated by both the resident and a nursing staff member. A resident was considered eligible for the study, if either she herself, or the staff member, or both came up with a total score ≥ 12).

The level of premorbid intelligence was globally assessed by the staff of the homes for the elderly and the investigators involved in the screening examination. Time constraints during the screening interview precluded the application of one of the more formal methods that have been developed to test this patient characteristic^{20,21}.

The absence of dementia was verified by means of the findings obtained from the SIDAM questionnaire (see previous heading: "Mild to moderate dementia").

Exclusion criteria

The following criteria precluded the inclusion of patients in our trial:

1. Age <50 years.
2. Exposure to a co-intervention with any drug or manoeuvre that was prescribed deliberately as a treatment for dementia or cognitive function impairment (e.g., nootropics) or to co-medication with any drug or chemical substance that was prescribed for other indications than dementia or cognitive impairment, and that may have a substantial (deteriorating) impact on the psychical or cognitive functioning (see appendix 1).
3. The presence of any medical condition or somatic factor that might be a non-treatable underlying cause of the demential symptoms or the cognitive com-

- plaints, or that might otherwise interfere with the trial conditions (see appendix 2).
4. The presence of a severe state of depression.
5. The presence of any absorption disorder that might limit the bio-availability of essential Ginkgo substances.
6. Lack of compliance demonstrated during the run-in phase.
7. Recovery of original cognitive impairment by the end of the run-in phase.
8. Expected drop-out before the end of the study due to, for instance, nursing home admission, long-term hospital stay, or mortality.
9. Contra-indications against EGb 761 medication.

The use of certain types of co-medication was not allowed at all during the trial. However, candidate participants got the opportunity to have these substances washed out during the run-in phase of the trial. Co-medication with other cognition-related pharmaca was allowed, provided that a 'steady state' could be assumed and that indications for interference with the trial medication were absent. Appendix 1 gives an detailed overview of the drugs concerned.

Any use of such medicines was documented, both before and during the trial. Although there were small differences in procedures between the homes for the elderly, a 'medicine sheet' was available for each resident during the trial, from which both the current and the past administration of pharmacotherapeutics could be extracted. Sometimes the administration of medicines was fully automatized, and arranged directly by a pharmacy ('pill packages'), without mediation of the caregiving personnel. Medicine supply along other channels than the regular one, and self-medication - for instance: food supplements with an anti-oxidative potential, like vitamin C, vitamin E, selenium or

beta-carotene - were also taken into account.

With regard to co-morbidity, we decided to deal with some conditions as criteria for exclusion in an absolute sense (e.g., Parkinson's disease), whereas other conditions only gave rise to exclusion when a sufficient amount of evidence for a causal relationship with the target disease of the trial appeared to be present. Appendix 2 contains a list of the diseases and conditions that might result in exclusion from the trial.

The presence of these conditions was assessed through medical history taking. The general practitioner and the head of the caregiving team in the old people's home were asked to screen a checklist of suspicious health conditions for each subject involved in the screening procedure. All identified conditions were documented and evaluated in order to decide on preclusion from the trial.

Medical screening was mainly based on pre-existing knowledge. CT-scans, Doppler sonography and other laboratory tests were not performed as part of the trial. However, if such examinations had been done in the past, serious efforts were made to retrieve the results.

Since a state of depression can be accompanied by temporary manifestations of cognitive dysfunction which may not be amenable to Ginkgo treatment, severely depressed persons had to be excluded from the study population. The setting of the trial - e.g., no expertise near at hand - did not favour the clinical assessment of affective disorders. Therefore, we tried to identify seriously depressed persons through a self-assessment scale: the GDS (Geriatric Depression Scale), consisting of items with a yes/no response format, was chosen to measure the state of depression. The single items reflect several aspects of (depressive) mood that have been noticed by the respondent during

the past week²². Previous research has shown that the GDS, if compared with clinical assessment procedures, results in a sufficiently valid estimation of the state of depression, especially for people who are not severely demented. Instead of the original GDS version, which consists of 30 items, we preferred to use a short 15-item version, resulting in a total score ranging from 0 (most favourable) to 15 points²³. Candidates with a score ≥ 11 , indicating severe feelings of depression²⁴, were excluded from the study population. During the course of the trial the GDS was scored three times, in order to detect a possible effect of Ginkgo on depression.

In addition, the presence of depressive mood was also assessed as part of the SCAG (see section 5.1.1).

Serious contra-indications against treatment with *Ginkgo biloba* special extract had never been reported. From earlier investigations it was known that, occasionally, an innocent allergic reaction may occur (dermatologic complaint: itching). Very rare cases of reversible, non-specific gastrointestinal complaints had also been reported. Based on previous experiences both types of adverse events were expected to occur in less than one percent of the Ginkgo users.

3.1.3 Sample size

The target number of people to be enrolled in the study was set to 300 (lower limit). Recruitment of this number of participants seemed to be feasible within a period of 12 to 18 months, given the estimated frequency of occurrence of dementia and related cognitive disorders in the open community, especially among the oldest old, and the even higher prevalence in institutions such as homes for the elderly and nursing homes^{25-29,30}.

Many outcome measures in this trial were based on a (semi-)continuous scale, which

makes the study relatively efficient from the viewpoint of sample size requirements. Sample size calculations were made in advance, although the complex structure of the trial (repeated random allocation; initially three, later five trial arms), and uncertainties about the distribution characteristics and the clinical meaningfulness of the outcome scales in this trial make it difficult to rely heavily on formal sample size and power calculation techniques. Focusing on the treatment effect for Ginkgo (elevated and usual dosages combined) compared to placebo after 24 weeks of intervention, we calculated, for instance, that in order to detect a 3-point difference for the SKT as the outcome measure of interest (scale 0-27; see section 5.1.1), a sample size of 66 - 22 for each of both Ginkgo groups and 22 for the placebo group - would be sufficient, assuming a population standard deviation for the SKT of 4 scale-points, and with an allocation ratio of 2 : 1, $\alpha = 0.05$, and $\beta = 0.20$. Likewise we found that, to detect a two-fold reduction of the proportion of participants who have been deteriorated under Ginkgo compared with placebo according to clinical global impression (CGI, see section 5.1.1), a total sample size of 150 would do, assuming a deterioration rate of 50% for the placebo group, and given an allocation ratio of 2 : 1, $\alpha = 0.05$, and $\beta = 0.20$. And with regard to a responder analysis, based on the three primary outcome measures (see section 5.1.1), a sample size of approximately 210 would be sufficient to identify a two-fold increase in response percentage under Ginkgo, assuming 20% responders in the placebo group (same treatment allocation ratio and acceptance levels of type I and II errors). Eventually we decided to aim at about 50 evaluable study participants in each of the five trial arms. Anticipating a 20% drop-out rate, we attempted to include 300 eligible subjects.

3.1.4 Screening procedure

Screening for dementia and AAMI

In each of the 39 participating homes for the elderly, the selection of eligible study subjects was performed through a number of consecutive steps:

1. As a first step the nursing staff of each home was asked to make a pre-selection of residents who seemed to be appropriate to participate in the initial screening examination. This provisional selection was based on pre-existing information on (observable) shortcomings in cognitive functioning.
2. Residents who had been recommended by the staff were then informed about the planned investigations by an individual letter. A similar letter was sent to a close relative. Through this letter both the resident and her relatives were invited to attend an informative meeting with the research team. At the end of this meeting they were asked whether they were interested in participating and whether they agreed with the retrieval of some medical background information.
3. Next, information was collected on the health status and the medicine use of the pre-selected subjects. This information - derived from the medical administration in the institution - was used to check the inclusion and exclusion criteria regarding medical conditions and drug consumption. At this stage the general practitioner was approached, to verify the absence of serious objections against participation on medical grounds and - if necessary - to provide additional medical background information.
4. Those residents who had qualified themselves for the next step were invited for the initial screening examination. Through an interview they were screened for:

- Presence / absence of mild or moderate dementia.

The test part of the SIDAM (introduction; questions 1-40, including the MMSE) was completed during this interview session, whereas the clinical assessment part (questions 41-52) was completed separately by the nursing staff.

- Presence / absence of impairment in memory and attention.

Objective testing by means of the Syndrom Kurz Test (SKT), in combination with subjective self- and hetero-assessment of memory problems in everyday life (modified MAC-Q).

- Level of premorbid intellectual state (estimated by one of the researchers and the staff).

- Presence / absence of a state of depression (short version of the GDS). Furthermore, during this interview we gathered or checked information regarding personal characteristics (e.g., name, age, gender, marital status, weight, and height).

In order to fulfil the informed consent requirements, the purpose and the design of the study as well as the trial conditions were explained in detail before the start of the interview.

5. After completing the initial screening examination, those people who met the inclusion criteria for mild to moderate dementia, memory impairment and intellectual capacity and who were (still) willing to participate, entered the run-in phase. During this 3-week period they received placebo treatment. In the meantime the information collected during the screening examination was processed further and the candidate participants were given the opportunity to reconsider their preliminary decision to participate.

6. At the end of the run-in phase for each study candidate the level of compliance with the (placebo) intervention was assessed. Those subjects who showed a sufficient level of compliance and a rather stable clinical picture with regard to memory function and other dementia-related symptoms, who still met the other inclusion and exclusion criteria, and who were prepared to sign the written informed consent form, qualified for the baseline measurement and the subsequent random allocation procedure. Unless the complaints reported at the time of the initial screening examination seemed to have faded away, they were randomly allocated to one of the treatment groups. This time-point demarcated the actual start of the intervention.

Informed consent procedure

Before the start of the study, as part of the initial screening examination, each study candidate received comprehensive information - both in oral and in written form - about the purpose, the design and the possible consequences of the trial: study question, characteristics of the compound under examination, supposed mechanisms of action of *Ginkgo biloba* extract, experiences from previous research and testing, possible side effects, treatment conditions and duration of the treatment period, double-blind and placebo-controlled approach, essentials of the randomization procedure, type and timing of the measurement procedures (initial screening examination, psychometric testing), opportunity to report side effects in the meantime, opportunity to withdraw from the trial prematurely (without reasons given and without negative consequences), opportunity to collect information on further details and on results of the study.

The written consent form was signed by each participant. In nearly all cases a close relative or friend was involved in the decision

making process and - if necessary - this person co-signed the consent form. Moreover, consent could only be agreed upon after the general practitioner in charge of the regular medical care for the study subject, had been given the opportunity to put forward any objections against the participation of his client.

The informed consent procedure was executed in agreement with the regulations of the national Dutch legislation and the local medical-ethical committee of the Academic Hospital Maastricht / Maastricht University. Definite consent for each eligible resident was established at the time of the baseline measurement session, after finishing the qualification period.

Results of the screening procedure

As stated before, a total number of 4,459 elderly people were residing in the cooperating 39 homes for the elderly. In 7 homes the selection procedure was performed twice, with an interval of at least one year in between. Approximately 800 residents (18% of the total population of residents) were preselected by the staff of the homes as potential candidates for the trial. About 700 of them fulfilled the first set of eligibility criteria (interest in participation, check of co-medication and co-morbidity). 513 Persons were actually screened (initial screening interview and staff assessment). Half of these residents (256, or 6% of the total population of residents) fulfilled the remaining selection criteria as well, and agreed to participate in the study. They entered the run-in period of 3 weeks. At the end of this period, another 41 residents had to be excluded, for various reasons (e.g., hospital admission in the mean time, problems experienced with the intake of the study medication, refusal to participate after all).

Eventually, 215 participants (5% of the total population of residents) were randomised to one of the three treatment groups that en-

tered the first intervention period of 12 weeks. Due to an administrative mistake, one woman was randomised erroneously: although she gave informed consent, she refused to participate in the baseline interview and did not receive any study medication. Therefore, the final study population consisted of 214 randomised participants. Appendix 3 gives a detailed overview of the selection procedure and the reasons for withdrawal during each phase of the procedure.

3.2 COMPARABILITY OF TRIAL SUBGROUPS

3.2.1 Prestratification and randomization

Admissible subjects were allocated at random to one of the treatment groups. Randomization was performed within separate strata for diagnostic subcategory (age-associated memory impairment or mild/moderate dementia) and speed of development of the memory complaints (perceived deterioration of the complaints during the last 6 months: yes/no). Thus, the participants were pre-stratified into 4 strata and subsequently randomly allocated to the treatment groups within each of these strata.

Instead of allocating the eligible subjects already at the start to one of the five resulting treatment groups, randomization was performed as a two-stage procedure, that means separately for each of the two intervention periods (week 1-12 and week 13-24). Immediately after baseline measurement the participants were allocated to one of the three treatment groups that would enter the first intervention period: EGb 761 usual dose, EGb 761 elevated dose, or placebo treatment. Originally the allocation ratio's between these treatment groups were 2 : 2 : 1. These were changed to 1 : 1 : 1 from January 1996 onwards, to enhance the

Table 1. Numbers of study participants randomized to the alternative treatment regimens.

Treatment group	1st intervention period	2nd intervention period
Placebo	n = 48	n = 122
Verum 160 mg	n = 84	n = 40
Verum 240 mg	n = 82	n = 39
Total	n = 214	n = 201

power of the trial with regard to the analysis of the effect associated with the main treatment contrast, between Ginkgo and placebo after 24 weeks of intervention. At the end of week 12, those individuals in the verum groups who were still participating, were randomly allocated to either continued treatment with EGb 761 or placebo (EGb 761 usual dose or placebo, ratio 1 : 1; and EGb 761 elevated dose or placebo, ratio also 1 : 1). Those participants who received placebo during the first intervention period, continued the use of placebo tablets during the second intervention period. This repeated randomization strategy was applied to prevent inefficiency. This might occur if the numbers of drop-outs would be unequally distributed among the treatment subgroups in the first intervention period.

For the concealed randomization procedure we used a computer generated random sampling set. The first randomization within each of the strata was performed in blocks of 5 consecutive participants, whereas the block size for the second randomization was set to 2. Copies of the randomization lists were sealed and locked up at two separate addresses of staff members of the Department of Epidemiology, who were not directly involved in the practical execution of the trial. Table 1 shows the number of participants in each treatment group resulting from the randomization procedure. 13 Subjects dropped out for various reasons during the first intervention period, and could not be randomized for the second period.

3.2.2 Confounders and effect modifiers

Factors known or suspected to influence the course of cognitive decline were considered potential confounders or modifiers of the envisioned effect of Ginkgo treatment in this trial. Several of such factors have been described in the literature³¹⁻³³. In this respect it should be noted that etiologic and prognostic determinants are difficult to distinguish, partly due to the insidious onset of the symptoms. Therefore, we considered both types of determinants as potential confounders or effect modifiers. Several of these variables were measured in our trial - as a rule only once, at the time of the initial screening examination - and taken into account by means of exclusion (see above), prestratification (see above), and/or multivariate analysis. Some events that may be associated with a relatively early and/or fast cognitive impairment during the aging process have been integrated into a checklist of so-called 'biological life events' (BLEs) by Houx et al³¹. From a cross-sectional study with volunteers recruited from the normal, healthy population enclosing a broad age range, it appeared that subjects with at least one BLE in their history accounted to a large extent for the age-associated decrease of scores on a battery of cognitive function tests³¹. At the time of onset of the Maastricht Ginkgo Trial the nine categories of BLEs that were distinguished were not yet incorporated into a structured index. It was suggested that these categories should be handled separately

and in a flexible way. Based on their estimated relevance and the opportunity to retrieve useful information, we selected the following life events for our study: history of brain damage (closed head injury, concussion, post traumatic amnesia); history of general anaesthesia; intake of alcoholic beverages; regular use of 'yellow sticker' drugs that affect one's driving ability or level of consciousness; and history of any neurological disorder and/or neurological treatment (e.g., meningitis, encephalitis, migraine), which had resulted in complete recovery. By means of a short semi-structured interview the presence of each individual BLE-item was evaluated. Moreover, the total number of BLEs was recorded.

Eventually, data on the following background variables were recorded:

- age
- gender
- body weight and body height, allowing for calculation of the Quetelet Index (QI)
- pre-treatment severity of cognitive dysfunction
- estimated speed of deterioration of memory complaints during the preceding 6 months
- use of co-interventions during treatment
- BLE-score.

4 STUDY TREATMENT

4.1 CHOICE OF INTERVENTION

4.1.1 Characteristics of the study treatment

Drug formulation

The verum preparation in our trial consisted of round film-coated tablets, 10 mm in diameter, containing either 80 mg or 120 mg of the dry extract EGb 761[®] from the leaves of the *Ginkgo biloba* tree (35-67:1 ratio of raw plant material to extract), adjusted to 24% Ginkgo flavonol glycosides and 6% terpene

trilactones (ginkgolide, bilobalide). Ginkgolic acids were kept below 5 ppm. The tablets were built up from a matrix consisting of lactose, silica, cellulose, maize starch, sodium carboxymethylcellulose, talc, and magnesium stearate, in which the brown extract was embedded. The yellow film coating consisted of hydroxypropyl methylcellulose, polyethylene glycol, titanium dioxide, iron (II) oxide, and carnauba wax.

The placebo preparation consisted also of film-coated tablets, which contained no *Ginkgo biloba* extract at all (0 mg). Aiming at a pure intervention contrast, much effort was paid to develop a placebo formula that matched the verum formula perfectly with regard to all relevant aspects: shape, size, surface colour and texture, weight, taste on licking and chewing, internal colour and structure, smell, specific gravity, and also packaging and labelling³⁴. Comparability with regard to shape, size, surface colouration and texture, weight, taste on licking, and - to an acceptable approximation - smell and specific gravity, were ensured by pressing placebo tablets of the same shape and size from the same matrix, and by covering them with an identical film coating. To ensure that the internal colour and texture also matched, caramel and brown iron oxide were added to the tablet matrix. This resulted in a perfect match in colour and internal texture, so that any distinction between the placebo formulation and the EGb 761[®] tablets could be ruled out, even after breaking or crushing. To mimic the pronounced bitter taste that is a characteristic of the Ginkgo extract, 2 mg of quinine hydrochloride were added to each tablet, after addition of a bitter principle derived from a gentian extract and of quinine sulfate to placebo test batches had failed to yield satisfying results. The bitter taste of these placebo tablets was found to be equally intense and unpleasant as that of the EGb 761[®] tablets, before the start of the trial

and in a further test 3½ years later, after the end of the study.

Breaking, crushing and grinding did not reveal any differences in physical characteristics between the Ginkgo and placebo tablets. However, repeated tests showed that the initially placebo tablets swelled and completely disintegrated within one minute of being immersed in water, whereas the EGb 761® tablets, owing to the lipophilic components of the Ginkgo extract, swelled slowly and only began to disintegrate after about 11 minutes. After addition of hydrogenated castor oil, a common excipient in tablet manufacture, the disintegration time of the placebo tablets was sufficiently prolonged to prevent any distinction from the EGb 761® tablets.

Caramel, brown iron oxide and hydrogenated castor oil were added only to the placebo tablets, not to the EGb 761® tablets. These substances can, however, be regarded as pharmacologically inactive. By contrast, quinine, which was added to the placebo tablets only, to avoid any risk of interactions with the active substance, is in no way pharmacologically inert. We believed, however, that the low dosage applied would preclude any pharmacological effect that is similar or counteracts to the Ginkgo effect. In the treatment of tropical malaria the recommended daily dose for adults is 1950 mg quinine sulfate, which is more than 450 times the daily dose in our trial (4 mg). Two placebo tablets contain even less quinine than half a glass of typical tonic water, a popular drink with a content of some 52 mg per litre³⁴.

Basic dosage schedule

The Ginkgo and the placebo tablets were administered orally in a fixed dose. The tablets had to be swallowed with water or another beverage, not bitten or crushed. During the run-in period each potential participant received 2 x 1 placebo tablet each day (morning and evening, during the meals). Subjects allocated to verum group 1 (160 mg of

Ginkgo per day) swallowed 2 times 1 tablet EGb 761 (80 mg) each day (morning and evening, during the meals) during the first intervention phase, and 2 times 1 tablet EGb 761 (80 mg) or 2 times 1 placebo tablet during the second intervention phase. Subjects allocated to verum group 2 (240 mg of Ginkgo per day) swallowed 2 times 1 tablet EGb 761 (120 mg) each day (morning and evening, during the meals) during the first intervention phase, and 2 times 1 tablet EGb 761 (120 mg) or 2 times 1 placebo tablet during the second intervention phase.

Those subjects allocated to the placebo group (0 mg of Ginkgo per day) received 2 times 1 placebo tablet each day (morning and evening, during the meals), during both the first and the second intervention phase.

Duration of therapy

After completing the qualification period successfully each eligible person was treated for 24 weeks (12 + 12: verum 1 + verum 1, verum 1 + placebo, verum 2 + verum 2, verum 2 + placebo, or placebo + placebo). So, by the end of the trial approximately 380 pills ((21 x 2) + (84 x 2) + (84 x 2)) should have been consumed by each cohort member.

Packaging and distribution of the tablets

The packaging and labelling of the preparations were in accordance with the requirements of the Dutch Law of Medicine Supply and the Good Clinical Practice guidelines. The tablets were packed in blisters, with each blister containing 10 tablets. The blisters in turn were packed in boxes, with each box containing 7 blisters, i.e. a little more tablets than the prescribed monthly dose for each participant. These boxes were supplied to each participating home on three occasions: start of the run-in period, start of week 1-12, and start of week 13-24. The tablets were distributed to the study subjects either on a daily basis or immediately preceding each meal, at least if the medicine supply

within the home was coordinated by the nursing staff itself. In some cases, if the provision of medicines had been boarded out to a local pharmacy that relied on an automated distribution system, the supply was organized on a weekly basis.

Each blister box was labelled with information regarding the batch number of the tablets, the period of tenability, the producer's name, the name of the principal investigator, instructions for the use of the medicines, the content and composition of the tablets, the number of tablets present within each package, the specification "for clinical research", the name of the participant, and a unique randomization number for each participant. Name, randomization number and guidelines for use of the medication were also printed on a separate label attached to each blister. The packaging and labelling characteristics were identical for the verum tablets and the placebo tablets.

4.1.2 Qualification period

A 3-week qualification period (run-in phase) was planned to precede the actual intervention. Such a period can be used to pursue various objectives, for instance individual titration of the treatment dose³⁵. In our trial the run-in period mainly served the following purposes:

- To 'wash-out' any residual effects of previous treatments.
- To verify the stability (or progress) of the target symptoms and to detect cases of spontaneous regression of relevant indicators of cognitive impairment. This allowed for the identification and exclusion of strong responders to placebo before the start of the trial.
- To identify (potential) non-compliers to the prescribed drug regimen.
- To add sense to the informed consent procedure and to the ethics of the trial in general.

As described previously in section 3.1.4, the run-in period turned out to be successful, since 41 study candidates (out of 256 included in this period) could be 'saved from randomization'. Probably a large part of them would have dropped out during one of the subsequent intervention periods, thereby threatening the internal validity of the trial.

4.1.3 Compliance

Poor compliance to the prescribed drug regimen might also threaten the validity of the trial. In this trial compliance was stimulated by means of careful instruction at the onset of the trial. One advantage of the setting of the trial was that it is common practice in Dutch homes for the elderly to control the use of any prescribed medication on a daily base. If necessary, residents are even asked to take their pills under direct supervision. In addition to this form of therapeutic control, pill taking behaviour was assessed by interview and regular pill counting. At each measurement point the experiences of the participants with the drug regimen were evaluated briefly.

4.2 COMPARABILITY OF INTERVENTION CONDITIONS

We believe that in our trial the use of a placebo intervention guaranteed a high degree of comparability of the intervention conditions. We observed a close similarity between the Ginkgo and the placebo preparations.

The thoroughness of the blinding strategy was assessed by asking the study participants, the nursing staff members, and the interviewers to give their opinion about the type of tablet used, both 4 and 18 weeks after the start of the trial.

In the section dealing with the patient selection procedure we described already which

co-medication and other types of co-intervention were not allowed at the start of the trial. Differences between the treatment groups regarding the exposure to co-interventions were registered during the intervention period, and accounted for in the analysis. Standardization of co-interventions was not pursued.

5 STUDY OUTCOMES

5.1 CHOICE OF OUTCOME MEASURES

5.1.1 Measurement of intended outcomes (efficacy)

To diagnose and follow the course of manifestations of cognitive impairment and dementia, various types of measurement tools can be used: medical history (reported by study subjects, close relatives, or medical practitioners), personal experience (self-assessment by study subjects), clinical assessment (hetero-assessment through observation and structured interviews), neuropsychological tests, laboratory assessment, etc. We decided to put emphasis on clinical assessment and neuropsychological testing of various aspects of cognitive functioning as central outcome measures. In addition, we paid attention to social functioning and functioning with regard to activities of daily living. In view of the setting of our trial we made no special efforts to try to enhance the accuracy of the diagnostic process by means of laboratory measurements, such as electrophysiological methods, computerized tomography (CT), and regional cerebral blood flow (rCBF)^{36,37}.

Several general considerations guided the selection of the outcome measures in our trial:

- To be feasible the baseline measurements and the follow-up measurements

had to be amenable to completion within a single session lasting 60 minutes at maximum.

- The neuropsychological function tests had to cover the most relevant aspects of cognitive function in general, and memory function in particular.
- Preferably, existing psychometric tests had to be applied, that had been validated already extensively, and that are supplied with several parallel forms, allowing for repeated measurements without introducing a test effect.
- The neuropsychological tests should be valid, precise, and sensitive enough to enable the detection of small but clinically relevant changes in cognitive function within a rather short time-span.
- The level of sophistication of the clinical assessment scales and the psychometric test battery should not preclude their application by a trained research assistant.

Figure 2 displays an overview of the measurement tools (tests, questionnaires, rating scales) that were used in the trial.

Initial screening examination

In general the initial screening examination (interview, assessment by nursing staff) took place three weeks before the baseline measurement, which marked the start of the intervention (timepoint -3 in figure 2). The following information was collected through this screening examination:

- Assessment of dementia: SIDAM (including MMSE and Ischemic Score).
- Assessment of memory impairment (objective) : Syndrom Kurz Test (SKT).
- Self-assessment of memory impairment (subjective): modified MAC-Q (5 items), stability of memory complaints over the previous six months.
- Self-assessment of depression: Geriatric Depression Scale (GDS, 15 items).

Measure	Time schedule (weeks)						
	-3	0	4	8	12	18	24
<i>Clinical assessment</i> (dementia, cognitive functioning, depression, geriatric health complaints, etc.):							
• Dementia, AAMI (SIDAM)	X						
• Sandoz Clinical Assessment-Geriatric Scale (SCAG; hetero-assessment)		X	X	X	X	X	X
• Clinical Global Impression (item 2; hetero-assessment)			X	X	X	X	X
• Depressive mood (GDS-15 items; self-assessment)	X				X		X
• Health status (report mark; self-assessment)		X	X	X	X	X	X
• Memory status (report mark; self-assessment)		X	X	X	X	X	X
• Occurrence of specific health complaints (dizziness, headache, etc.; hetero-assessment)		X			X		X
• Memory complaints (modified MAC-Q; self-assessment)	X						
• Memory complaints (modified MAC-Q; hetero-assessment)	X						
• Stability of memory complaints	X						
<i>Neuropsychological assessment</i>							
• Syndrom Kurz Test (SKT; 9 subtests)	X		X	X	X	X	X
• NAI-WL (Wort Liste: verbal learning test)		X			X		X
• NAI-ZN (Zahlen Nachsprechen: digit span)		X			X		X
• NAI-ZVT (Zahlen Verbindungs Test: trail making test)		X			X		X
<i>Behavioural assessment</i>							
• NAI-NAA (Nürnberger Alters Inventar - instrumental activities of daily life of elderly people; hetero-assessment)		X			X		X
• NAI-NAA (Nürnberger Alters Inventar - instrumental activities of daily life of elderly people; self-assessment)		X			X		X
<i>Background characteristics</i>							
• Demographic characteristics (age, sex, etc.)	X						
• Biological Life Events (BLEs)		X					
• Psychological Life Events (PLEs)	X	X	X	X	X	X	X
• Level of premorbid intelligence	X						
<i>Other phenomena of interest</i>							
• Check of inclusion / exclusion criteria	X	X					
• Informed consent		X					
• Compliance with study medication		X	X	X	X	X	X
• Adverse events		X	X	X	X	X	X
• Addiction / abstinence phenomena						X	X
• Change in memory complaints in comparison with screening		X					
• Check of blinding			X			X	
• Medical status: disease history	X						
• Medical status: medication history	X						
• Exposure to co-interventions		X			X		X

Figure 2. Timing of baseline and outcome measures.

- Background variables: age, sex, height, weight, intelligence level, self-medication.

In addition, information on the presence of relevant diseases and drug therapy was obtained from the staff of the homes.

Outcome measures: baseline and follow-up examinations

In accordance with the most recent standards for the evaluation of anti-dementia drugs and cognition enhancers, outcome parameters were defined at three distinct levels of functioning³⁸⁻⁴⁰:

- Cognitive functioning.
Performance measured by neuropsychological function tests.
- (Psycho-)pathology.
Clinical signs and symptoms. Clinical assessment was pursued by both self-assessment (subjective status) and hetero-assessment, both at a global and at a more specific level.
- Behavioural functioning.
Functional capacity with regard to (instrumental) activities of daily life which are considered appropriate for people residing in homes for the elderly. Behavioural assessment was pursued both by self-assessment and by hetero-assessment (behavioural observation).

Furthermore, a distinction was made between outcome measures of primary (confirmatory analysis) and secondary (descriptive analysis) interest. Outcome parameters were assessed at several consecutive time-points. After the baseline measurement at week 0 (day 0), follow-up measurements were performed at week 4 (day 28), week 8 (day 56), week 12 (day 84), week 18 (day 126), and week 24 (day 168).

Primary outcome measures

The main outcome parameters covered the three domains described above. The following measures were chosen to represent the principal study outcomes:

1. Syndrom Kurz Test (SKT).

This psychometric test battery, consisting of nine subtests, is a quick test to measure memory and concentration impairment. Details of this test have already been described in section 3.1.2 of this paper. Although five parallel forms (A-E) were available for repeated measurements, we used only three (A-C) in this trial, in an alternating sequence. The overall test score ranges from 0 to 27 points.

2. Clinical Global Impression, item 2 (CGI-2).

This hetero-assessment scale is a popular tool to register a global overall impression of any clinical changes observed. The CGI contains three items, which relate to the degree of severity of disease, the change in status, and an index of therapeutical efficacy.

The second CGI-item, which was used in this trial, stands for a direct comparison of the current relevant health state of the study subject - in this case: level of cognitive functioning - and the health state at the start of the intervention period, without attributing this change to the action of the medication (7-point scale: very much improved, much improved, a little bit improved, unchanged, a little bit deteriorated, much deteriorated, very much deteriorated). In our study the assessment was made by the nursing staff.

3. Nürnberger Alters Inventar, Nürnberger-Alters-Alltagsaktivitäten-Skala (NAI-NAA).

This self-assessment scale has been designed to measure the independent functioning of elderly people regarding daily activities. The NAA consists of 20 items related to

instrumental activities (e.g., getting dressed, performing light household tasks), social activities (e.g., visiting other people, participating in social events), and general cognitive performance (writing letters or postcards, retrieval of personal belongings). Each item has to be scored using a 3-point format ("I do this often, sometimes, or never"). The total test score ranges from 20 to 60 points⁴¹. In our trial the NAI-NAA was filled out both by the residents themselves and by the nursing staff (hetero-assessment). The nursing staff assessment was used as one of the primary outcome measures.

Secondary outcome measures

In addition to the main outcome parameters the following measurement tools were used to complete the picture:

1. Nürnberger Alters Inventar, Zahlen-Verbindungs-Test-G (NAI-ZVT-G).

The ZVT-G is an adapted form of the 'Trail Making Test'⁴²; it is a paper-and-pencil task implying that 30 numbers (1-30) have to be connected with each other in the right sequence and as quickly as possible. It measures aspects of planning, organization and flexibility. The test has proven to be able to differentiate between people with and without organic brain syndrome⁴¹.

2. Nürnberger Alters Inventar, Zahlen-Nachsprechen-G (NAI-ZN-G).

The ZN-G ('Digit Span Test') is an adapted form of the 'digit span' test which originates from the WAIS and the (revised) Wechsler Memory Scale (WMS)^{43,44}. It intends to quantify the size of the short term memory, but it also relates to the complex operations involved in the organization of information. The test is appropriate to distinguish further between people with a rather low level of intellectual functioning. The task implies that series of digits of increasing length, read by the test leader, are reproduced immediately

in the same sequence (forward). After finishing the forward test the procedure has to be repeated in a backward format (reproduction in reverse sequence) with series of digits of increasing length. The test score is calculated as the sum of the two digit-lengths that represent the highest correct achievement in forward and backward reproduction (range: 0-17)⁴¹.

3. Nürnberger Alters Inventar, Wortliste (NAI-WL).

The WL ('Verbal Learning Test') is also an adjustment of a procedure that previously has been introduced in several formats. Examples are the 'Vijftien Woorden-Test' (15-WT)⁴⁵, and the 'California Verbal Learning Test', which is conceived as a 'shopping list' and contains 16 words from four semantic categories. The NAI word list consists of 8 di-syllabic words, which have been selected carefully, based on their meaning, level of abstractness and frequency of occurrence. The words are read loudly, one word per two seconds. Immediately thereafter the test subject is invited to reproduce as many words as possible (free recall; only one trial). The number of correct recalls is registered (test score; range: 0-8). This part of the test is followed about 20 minutes later by a second one, in which the same words are read again, but now integrated within a 16-item list which also contains 8 distracter items (recognition). The test score is calculated as the number of correctly recognized minus the number of incorrectly recognized verbal items (range: 0-8)⁴¹.

4. Sandoz Clinical Assessment-Geriatric Scale (SCAG).

The SCAG was filled in by the nursing staff. At each measurement time point an assessment was made of the level of presence of several important psychogeriatric symptoms. The SCAG has been developed as a comprehensive, general-purpose rating

scale, covering 18 symptom areas, that are rated according to a 7-point format (not present, very mild, mild, mild to moderate, moderate, moderately severe, severe)^{46,47}. It concerns symptoms such as confusion, mental alertness, irritability, hostility, indifference to surroundings, emotional lability, fatigue, and dizziness. Added is an overall impression of the study subject (item 19). Together the 18 target symptom categories are supposed to cover four broad empirical areas of impairment: cognitive function, mood and behaviour, ability to cope with the ordinary activities of everyday living, and certain somatic complaints (fatigue, dizziness, poor appetite). In formal validation studies factor analysis has revealed a structure with five separate factors: cognitive dysfunction, interpersonal relationships, affective disorders, apathy, and somatic functioning.

5. *Self-assessment of health status.*

At each measurement time point the participating elderly were asked to give a report mark for their health status in general. This mark ranged from 0 (very bad) to 10 points (excellent). The subjects were supposed to be familiar with such a scale, as it corresponds with the usual rating system in Dutch schools.

6. *Self-assessment of memory.*

At each measurement time point the participating elderly were asked to give a report mark for their perceived memory as well. This mark ranged also from 0 (very bad) to 10 points (excellent).

7. *Geriatric Depression Scale (GDS).*

The GDS scale for the self-assessment of depressive mood has already been described in section 3.1.2.

8. *Nürberger Alters Inventar, Nürberger-Alters-Alltagsaktivitäten-Skala (NAI-NAA).*

This regards the self-assessment version of the NAI-NAA, which we described already as one of the primary outcome measures.

5.1.2 Measurement of unintended outcomes (adverse events; safety)

Information about adverse events occurring during the trial was gathered and documented systematically. This information was collected by means of a structured questionnaire at each measurement time-point. This health symptom checklist referred to several frequently occurring health complaints (headache, stomach pain, diarrhoea, nausea, dizziness, constipation, skin irritation, heart fluttering, sleeplessness, nervousness). In addition, (severe) adverse events were reported spontaneously by the nursing staff in between.

Incidentally, it has been suggested that the use of nootropica like *Ginkgo biloba* extract might result in a mild form of addiction, with withdrawal effects after stopping treatment. The current study design, especially the second part of the two-phase architecture of the trial, was pre-eminently suited to investigate this claim. Therefore, the study subjects were confronted with a short list of abstinence complaints connected with well-known forms of addiction, both halfway and at the end of the second intervention period (vomiting, fatigue, shaking, muscle pain, lack of concentration, irritation, transpiring). In accordance with the GCP guidelines the occurrence of serious adverse events during the trial, such as the death of a study subject or hospital admission, was documented and evaluated. By doing this, for each serious event the probability of any relation with the study medication was assessed.

5.1.3 Data collection

The screening interviews were conducted by three members of the research team (HS, EvR, MvD). Assistance was given by a research assistant of the department of Epidemiology and by two student trainees, who executed part of the screening interviews. In addition, screening data were collected through assessments by the nursing staff of the homes for the elderly. In general the assessment forms were filled by the head of the team that was responsible for the care for the participants residing at a particular department of the old people's home.

The baseline and follow-up interviews were conducted by a team of 9 skilled interviewers, all females, who were appointed by that time to the Data Entry Service of the Faculty of Health Sciences / Faculty of Medicine of the Maastricht University, and who had gained much experience in interviewing techniques in general, amongst others through training courses. To be prepared to apply the measurement tools in the current trial in a proper way, these interviewers received a special training organized by the research team in advance of the onset of the trial. Occasionally, a refreshing session was organized for the interviewers to exchange experiences and to discuss problems encountered during the field work. In the meantime individual support and feedback were given to the interviewers.

The baseline and follow-up staff assessments were conducted by the same nursing staff members who had also rated the screening forms. This process was guided and supervised closely by the research team (in particular HS).

In general repeated measurements of a particular study subject were conducted by the same rater (nursing staff member) and the same interviewer. Only in case of force majeure this principle was violated. Both the raters and the interviewers were blinded for the

treatment given. Follow-up measurements had to take place at least within 7 days from the date of appointment. Later on we decided to liberalize this criterion to a range of 10 days, in view of the per-protocol analysis.

5.2 COMPARABILITY OF OUTCOME MEASUREMENTS

The outcome phenomena were assessed under blinded conditions. Blinded outcome measurement was promoted by launching a placebo intervention. The thoroughness of blinding was checked at 4 weeks and 18 weeks after the start of the intervention.

6 DATA MANAGEMENT AND ANALYSIS

6.1 DOCUMENTATION OF MEASUREMENT RESULTS

Case Report Forms (CRFs) were kept to document all relevant information collected for each individual study participant. This information included the signed patient consent form and all data regarding the screening procedure and the outcome measurements. Moreover, the patient's personal dossier included information on, for example, identifying and background variables, medical and pharmacotherapeutical history, concomitant events and processes (diseases, medicine use, medical and non-medical interventions), and adverse events (type, date of onset, duration, intensity, consequences, etc.).

All data were collected on paper (questionnaires, special forms, etc.), provided with a header to record identifying information (randomization number, date of measurement, etc.). If necessary, corrections were made in accordance with the GCP prescriptions. These implied striking out the incorrect information, entering the correct information alongside the misreported information, and

marking the freshly entered information with the author's signature and the current date. The completed forms and questionnaires were gathered at the project office as soon as possible, and immediately checked by the project staff. After this check the data were entered in computer files and stored on electromagnetic disk. Back-up files were produced at regular intervals. All documents containing individual information (paper as well as computerized versions) were stored in safeguarded and locked up filing-cabinets.

The data collection and storage were inspected by the monitor of the trial (AK) on a two-weekly basis. During the course of the trial a quality audit was made by a independent and qualified authority (GQS - Gesellschaft für Qualitätssicherung in der klinischen Forschung GmbH, established in Dortmund, Germany). This procedure led to a positive audit certificate⁴⁸.

6.2 ANALYSIS OF STUDY OUTCOMES

6.2.1 Confirmatory analysis

The working hypothesis underlying this clinical trial was that treatment with EGb 761 produces stronger effects than treatment with placebo tablets in elderly people with age-associated memory impairment or mild / moderate dementia.

In the context of the confirmatory analysis the efficacy of the intervention was evaluated on the basis of the scores for the primary outcome measures: the difference in SKT-scores between the start (week 0) and the end (week 24) of the entire treatment period; the score for CGI-item 2 at the end of the entire treatment period (week 24); and the difference in NAA_{staff}-scores between the start (week 0) and the end (week 24) of the entire treatment period. The comparisons had to be made between the combined verum groups (normal dosage of EGb 761 +

elevated dosage of EGb 761) and the placebo group. We tested the following hypotheses.

On top of the hypothesis system we introduced a global null-hypothesis, H_{00} , stating that there is no difference in intervention outcomes between subjects treated with EGb 761 and subjects treated with placebo, for either of the three criteria for efficacy (SKT, CGI-2, NAA):

$$H_{00} = H_{01}^{two} \cap H_{02}^{two} \cap H_{03}^{two}$$

where:

H₀₁: The expected difference in SKT-score between the start of the treatment and after 24 weeks of therapy is the same for subjects who have received EGb treatment and for subjects who have received placebo treatment (two-sided test).

H₀₂: The improvement of the clinical picture (CGI-item 2) after 24 weeks of treatment is the same for subjects who have received EGb treatment and for subjects who have received placebo treatment (two-sided test).

H₀₃: The expected difference in NAA_{staff} score between the start of the treatment and after 24 weeks of therapy is the same for subjects who have received EGb treatment and for subjects who have received placebo treatment (two-sided test).

In addition we decided that, if this global null-hypothesis would be rejected - and, therefore, would suggest the presence of a treatment effect - the single hypotheses referring to each of the primary outcome measures would be tested stepwise to identify the specific parameter(s) leading to this treatment difference. Or, in terms of alternative hypotheses:

H₁₁: The expected difference in SKT-score between the start of the treatment and after 24 weeks of therapy is larger for subjects who have received EGb 761 treatment than

for subjects who have received placebo treatment (one-sided test).

H₁₂: The improvement of the clinical picture (CGI-item 2) after 24 weeks of treatment is stochastically larger for subjects who have received EGb 761 treatment than for subjects who have received placebo treatment (one-sided test).

H₁₃: The expected difference in NAA-score between the start of the treatment and after 24 weeks of therapy is larger for subjects who have received EGb 761 treatment than for subjects who have received placebo treatment (one-sided test).

These single hypotheses were arranged in a hierarchical order on the basis of subject-matter considerations. This hierarchy - 1. SKT 2. CGI 3. NAA - corresponds with the sequence in which they were presented. The argument behind this hierarchy is that it can be assumed that any improvement will manifest itself first of all in the neuropsychological test results, thereafter in the clinical assessment by the care-provider, and finally in the everyday behaviour of the study subject. Actually, this hierarchy implied that the test procedure would be terminated whenever one of the single null-hypotheses would be accepted. In that case, the null-hypotheses referring to the remaining outcome measure(s) would be accepted as well.

To test the global null-hypothesis the multivariate OLS-test (one-sided) as described by Kieser et al.⁴⁹ was used. The results of this test indicate whether there is an 'overall' EGb 761 effect, without specifying which outcome measure(s) can be held responsible for the suggested effect.

We decided in advance that the hypotheses regarding the SKT and the NAA would be tested by means of a two-way ANOVA procedure, with the factors treatment and centre (home for the elderly) as independent vari-

ables. Because of the large number of old people's homes that eventually became involved, we decided to replace this test procedure by Student's t-test, including treatment category (combined Ginkgo groups vs placebo) as the single independent variable, and to deal with the factor 'study centre' as an independent variable in a multivariate analytical procedure.

In addition, a multiple linear regression analysis of the changes in SKT- and NAA-scores was conducted. In order to control for confounding and to elucidate effect modification, the following variables were included as covariates in the regression model initially: age, gender, baseline value for the effect measure concerned (either SKT- or NAA-score), diagnostic subgroup (memory impairment, dementia), speed of deterioration of cognitive function before the start of the study, old people's home, and level of co-interventions.

It was planned in advance to test the hypothesis regarding the CGI score by means of the Mann-Whitney U-Test, stratified for centre. Again, due to the large number of old people's homes involved we were forced to skip the stratification procedure and to deal with the factor 'study centre' as an independent variable through a multivariate analytical procedure.

Subsequently, a polychotomous logistic regression analysis was performed to estimate the effect of study treatment (Ginkgo vs. placebo) on 24 week CGI-score, adjusted for the influences of several potential confounding factors. For this purpose the CGI (Item 2), which originally was expressed on a 7 point scale, was categorized into three groups, corresponding with perceived improvement (a little bit, much, or very much), stability (no change), or deterioration (a little bit, much, or very much) of the condition. In this way a trichotomous outcome variable was created, which could be handled as the dependent variable in a polychotomous logistic regres-

sion analysis. The following covariates were included in the model as independent variables: age, gender, diagnostic subgroup (memory impairment, dementia), speed of deterioration of cognitive function before the start of the intervention, old people's home, and level of co-interventions. Testing was done at a significance level of $\alpha = 0.05$ (one-sided).

In addition, the Ginkgo treatment effect was evaluated by means of a responder analysis. At the time of the start of the Maastricht Ginkgo Trial it was recommended for drug trials (e.g., through the guidelines for GCP) to state in advance the main outcome measures and the expected treatment response (improvement) for each outcome measure separately^{4,40}. We considered the following responses to be of major clinical relevance:

- CGI-2: State of subject: 'somewhat improved', 'much improved', or 'very much improved'
- SKT: Improvement of total score with at least 3 points (range of scores: -27 to +27)
- NAA: Improvement of total score with at least 4 points (range of scores: -40 to +40)

Furthermore, we defined a responder as a study participant showing clinically meaningful improvement for at least two of the three primary outcome parameters. The proportion of responders was calculated for each of the treatment groups, both after 24 weeks (5 groups) and after 12 weeks (3 groups) of intervention. Differences in response percentage between the study subgroups were tested with a Chi-square test (one-sided; $\alpha = 0.05$).

6.2.2 Descriptive analysis

In addition to the confirmatory analysis, we dealt with the other study questions through a series of descriptive analyses of the data at

hand. These descriptive analyses covered the following comparisons (figure 1 specifies the treatment groups 1a, 1b, 2a, and 2b):

- The combined verum groups 1a + 2a (usual dosage of EGb 761 + elevated dosage of EGb 761) were compared with the placebo group after 24 weeks with regard to the secondary outcome measures: self-assessment of activities of daily life (NAI-NAA_{self}), geriatric symptom list (SCAG), self-assessment of health status, self-assessment of memory, state of depression (GDS), trail making speed (NAI-ZVT-G), digit span (NAI-ZN-G), word listing (NAI-WL)
- The combined verum groups 1a + 1b + 2a + 2b (usual dosage of EGb 761 + elevated dosage of EGb 761) were compared with the placebo group after 12 weeks with regard to the primary outcome measures (SKT, CGI-2, NAI-NAA_{staff})
- The combined verum groups 1a + 1b + 2a + 2b (usual dosage of EGb 761 + elevated dosage of EGb 761) were compared with the placebo group after 12 weeks with regard to the secondary outcome measures (NAI-NAA_{self}, SCAG, self-assessment of health status, self-assessment of memory, state of depression (GDS), NAI-ZVT-G, NAI-ZN-G, NAI-WL)
- In order to unmask a possible dose-effect relationship, the verum group 2a (elevated dosage of EGb 761) was compared with the verum group 1a (usual dosage of EGb 761), both after 24 weeks and after 12 weeks, with regard to the primary outcome measures (SKT, CGI-2, NAI-NAA_{staff})
- Also in order to reveal a possible dose-effect relationship, the verum group 2a (elevated dosage of EGb 761) was compared with the verum group 1a (usual dosage of EGb 761), both after 24 weeks and after 12 weeks, with regard to the secondary outcome parameters (NAI-

- NAA_{self}, SCAG, self-assessment of health status, self-assessment of memory, state of depression (GDS), NAI-ZVT-G, NAI-ZN-G, NAI-WL)
- In order to study the persistency of any intervention effect, the combined verum groups 1a + 2a (usual dosage of EGb 761 + elevated dosage of EGb 761; entire intervention period) were compared with the combined verum groups 1b + 2b (usual dosage of EGb 761 + elevated dosage of EGb 761; half of the intervention period) after 24 weeks with regard to the primary outcome measures (SKT, CGI-2, NAI-NAA_{staff})
- Also in order to study the persistency of any intervention effect the combined verum groups 1a + 2a (usual dosage of EGb 761 + elevated dosage of EGb 761; entire intervention period) were compared with the combined verum groups 1b + 2b (usual dosage of EGb 761 + elevated dosage of EGb 761; half of the intervention period) after 24 weeks with regard to the secondary outcome measures (NAI-NAA_{self}, SCAG, self-assessment of health status, self-assessment of memory, state of depression (GDS), NAI-ZVT-G, NAI-ZN-G, NAI-WL).

The majority of the statistical analyses were performed with SPSS for Windows, Release 7.5⁵⁰. For the polychotomous regression analysis to estimate the adjusted effect of Ginkgo treatment on the clinical global impression (confirmatory analysis) we used BMDP Statistical Software⁵¹. Cuzick's test for trend - to test the probability of a dose-effect relationship - was conducted with Arcus Quickstat⁵².

6.3 HANDLING OF MISSING DATA

Missing values for any of the primary and secondary outcome variables were replaced according to the following algorithm:

1. Replacement according to the 'last observation carried forward' principle. As a first step missing data were replaced by the last valid value obtained for the same variable in the same study subject. For multi-item or multi-subtest variables this replacement procedure was conducted at the item / subtest level. This replacement strategy was only applied to missing values preceded by at least one valid follow-up measurement (so, after at least four weeks of intervention). The following sequence was used for the 'last observation carried forward' replacement procedure: replacement of 8 week follow-up missing data, 12 week follow-up missing data, 18 week follow-up missing data, and 24 week follow-up missing data, subsequently. For example: if a person had missing scores for week 8 and 18, the score for week 8 was replaced by the score for week 4 and the score for week 12 was used as the result for week 18.
2. Replacement according to the 'mean value of valid subtests' principle. As a second step, which was only applied to multi-item or multi-subtest variables, any remaining missing values, related to one of the relevant measurement time-points (screening, baseline, 4, 8, 12, 18, 24 weeks of follow-up), were replaced by the mean value calculated from the valid item or subtest scores obtained for the same study subject at the same measurement time point. This 'mean value of valid subtests' replacement strategy was executed provided that missing values were present for no more than 25% of the component items or subtests of a particular scale.
3. If any missing values remained after the first and the second replacement step, these were not replaced at all.

6.4 PROTOCOL DEVIATIONS: INTENTION-TO-TREAT ANALYSIS VERSUS PER-PROTOCOL ANALYSIS

In view of the analysis of trial results the following characteristics were considered serious protocol violations:

- Insufficient level of compliance with the study medication.
- Insufficient level of participation in the baseline and follow-up measurements, e.g., due to premature drop-out.
- Unacceptable level of exposure to co-interventions.

Withdrawals (drop-outs) from the study were documented in accordance with the GCP guidelines. Withdrawal occurred, for instance, when a study subject deceased or repealed the written consent that was signed earlier, or when serious intercurrent diseases or medical interventions precluded further participation.

All randomized study subjects were considered eligible for the intention-to-treat analysis, provided that valid or substituted data on the outcome variable(s) involved were available, both for the relevant follow-up measurement and for the baseline (reference) measurement, allowing for the calculation of change scores.

In order to be eligible for the per-protocol analysis a randomized study subject should satisfy the following set of criteria:

1. Acceptable level of compliance with study medication:
 - An estimated consumption of at least 75% of the number of pills prescribed for both the first (0-12 w.) and the second (13-24 w.) intervention period justified inclusion in the 24 week per-protocol analysis.
 - An estimated consumption of at least 75% of the number of pills prescribed

for the first (0-12 w.) intervention period justified inclusion in the 12 week per-protocol analysis.

2. Acceptable level of participation in the follow-up measurements:
 - For inclusion in the 24 week per-protocol analysis it was required that the study subject had contributed to both the week 24 interview and the week 24 staff assessment. Moreover, the actual dates of the interview should not outrange the dates planned in accordance with the study protocol with more than 10 days.
 - For inclusion in the 12 week per protocol analysis it was required that the study subject had contributed to both the week 12 interview and the week 12 staff assessment. Moreover, the actual dates of the interview should not outrange the dates planned in accordance with the study protocol with more than 10 days.
3. Acceptable level of co-interventions:
 - To justify inclusion in the 24 week per-protocol analysis any type of co-medication started or adjusted during either the first (0-12 w.) or the second (13-24 w.) intervention period, should not have influenced the cognitive functioning and behaviour of the study subject in a substantial way. This requirement has been assessed for each study participant separately.
 - To justify inclusion in the 12 week per-protocol analysis any type of co-medication started or adjusted during the first (0-12 w.) intervention period, should not have influenced the cognitive functioning and behaviour of the study subject in a substantial way. This requirement has been assessed for each study participant separately.

REFERENCES

- Kleijnen J, Knipschild P. Ginkgo biloba for cerebral insufficiency. *Br J Clin Pharmacol* 1992; 34: 352-8.
- Commission of the European Communities. *Good clinical practice for trials on medicinal products in the European community. Draft report*. Brussels: Commission of the European Communities. Director-General for Internal Market and Industrial Affairs, 1989.
- Keppel Hesselink J. 'Good clinical practice': noodzaak of mode? *Ned Tijdschr Geneesk* 1989; 133: 2108-11.
- GCP-voorbereidingscommissie. *"Good Clinical Practice" voor het onderzoek met geneesmiddelen in de Europese Gemeenschap*. Utrecht: Nederlandse Stichting ter bevordering van Medisch-Farmaceutische Research, 1993.
- APA. *DSM-III-R. Diagnostic and Statistical Manual of Mental Disorders*; 3th ed. Washington, D.C.: American Psychiatric Association, 1987.
- WHO. *ICD-10: Tenth revision of the International Classification of Diseases*. Chapter V (F): Mental and behavioural disorders (including disorders of psychological development). Clinical descriptions and diagnostic guidelines. Geneva: World Health Organization, 1991.
- Zaudig M, Mittelhammer J, Hiller W, et al. SIDAM - A structured interview for the diagnosis of dementia of the Alzheimer type, multi-infarct dementia and dementias of other aetiology according to ICD-10 and DSM-III-R. *Psychol Med* 1991; 21: 225-36.
- Zaudig M. A new systematic method of measurement and diagnosis of "mild cognitive impairment" and dementia according to ICD-10 and DSM-III-R criteria. *Int Psychogeriatr* 1992; 4 (Suppl. 2): 203-19.
- Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12: 189-98.
- Hachinski VC, Illif LD, Zilkha E. Cerebral blood flow in dementia. *Arch Neurol* 1975; 32: 632-7.
- Rosen WG, Terry RD, Fuld PA, Katzman R, Peck A. Pathological verification of ischemic score in differentiation of dementias. *Ann Neurol* 1980; 7: 486-8.
- Erzigkeit H. *SKT. Ein Kurztest zur Erfassung von Gedächtnis- und Aufmerksamkeitsstörungen. SKT-Manual*; 4th ed. Castrop-Rauxel: Geromed GmbH, 1989.
- Erzigkeit H. The SKT - a short cognitive performance test as an instrument for the assessment of clinical efficacy of cognition enhancers. In: Bergener M, Reisberg D, eds. *Diagnosis and treatment of senile dementia*. Berlin: Springer, 1989.
- Overall JE, Schaltenbrand R. The SKT neuropsychological test battery. *J Geriatr Psychiatry Neurol* 1992; 5: 220-7.
- Ihl R, Frölich L, Dierks T, Martin EM, Maurer K. Differential validity of psychometric tests in dementia of the Alzheimer type. *Psychiatry Res* 1992; 44: 93-106.
- Kim YS, Nibbelink DW, Overall JE. Factor structure and scoring of the SKT test battery. *J Clin Psychol* 1993; 49: 61-71.
- Lehfeld H, Rudinger G, Rietz C, et al. Evidence of the cross-cultural stability of the factor structure of the SKT short test for assessing deficits of memory and attention. *Int Psychogeriatr* 1997; 9: 139-53.
- Crook TH, Larrabee GJ. Age-associated memory impairment: diagnostic criteria and treatment strategies. *Psychopharmacol Bull* 1988; 24: 509-14.
- Crook TH, Feher EP, Larrabee GJ. Assessment of memory complaint in age-associated memory impairment: the MAC-Q. *Int Psychogeriatr* 1992; 4: 165-76.
- Schmand B. *Nederlandse Leestest voor Volwassenen (NLV)*. Manual. Lisse: Swets & Zeitlinger, 1992.
- Lehrl S, Triebig G, Fischer B. Multiple choice vocabulary test MWT as a validated short test to estimate premorbid intelligence. *Acta Neurol Scand* 1995; 91: 335-45.
- Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression

- screening scale: a preliminary report. *J Psychiatr Res* 1982; 17: 37-49.
23. Sheik JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. In: Brink TL, ed. *Clinical gerontology: a guide to assessment and intervention*. New York: Hawthorn Press, 1986: 165-73.
 24. Baker FM, Miller CL. Screening a skilled nursing home population for depression. *J Geriatr Psychiatry Neurol* 1991; 4: 218-21.
 25. Skoog I, Nilsson L, Palmertz B, Andreasson L-A, Svanborg A. A population-based study of dementia in 85-years-old. *N Engl J Med* 1993; 328: 153-8.
 26. Rocca WA, Hofman A, Brayne C, et al. Frequency and distribution of Alzheimer's disease in Europe: a collaborative study of 1980-1990 prevalence findings. The EURODEM-Prevalence Research Group. *Ann Neurol* 1991; 30: 381-90.
 27. Letenneur L, Jacqmin H, Commenges D, Barberger-Gateau P, Dartigues JF, Salamon R. Cerebral and functional aging: first results on prevalence and incidence of the Paquid cohort. *Methods Inf Med* 1993; 32: 249-51.
 28. Morgan K, Lilley JM, Arie T, Byrne EJ, Jones R, Waite J. Incidence of dementia in a representative British sample. *Br J Psychiatry* 1993; 163: 467-70.
 29. Hofman A, Rocca WA, Brayne C, et al. The prevalence of dementia in Europe: a collaborative study of 1980-1990 findings. Eurodem Prevalence Research Group. *Int J Epidemiol* 1991; 20: 736-48.
 30. Boersma F, Eefsting JA, Van den Brink W, Van Tilburg W. Prevalentie en ernst van het dementie-syndroom in bejaardenoorden en verpleeghuizen. Een onderzoek bij ouderen afkomstig van het platteland rondom Zwolle. *Tijdschr Gerontol Geriatr* 1995; 26: 57-64.
 31. Houx PJ. *Cognitive aging and health-related factors*. PhD Thesis. Maastricht: Rijksuniversiteit Limburg, 1991.
 32. Colsher PL, Wallace RB. Epidemiologic considerations in studies of cognitive function in the elderly: methodology and nondementing acquired dysfunction. *Epidemiol Rev* 1991; 13: 1-27.
 33. Larson EB, Kukull WA, Katzman RL. Cognitive impairment: dementia and Alzheimer's disease. *Annu Rev Public Health* 1992; 10: 431-49.
 34. Knipschild PG, Hörr R, Oschmann R, Van Rossum E, Van Dongen MCJM. Optimization of placebos for double-blind clinical trials: experience with a phytopharmaceutical. *Arzneim.-Forsch./Drug Res*. 1998; 48: 1033-6.
 35. Knipschild P, Leffers P, Feinstein AR. The qualification period. *J Clin Epidemiol* 1991; 44: 461-4.
 36. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease; report of the NINCDS-ADRDA workgroup under the auspices of Department of Health and Human Services task force on Alzheimer's Disease. *Neurology* 1984; 34: 939-44.
 37. Scheltens P, Weinstein HC, Van Gool WA. De waarde van de nieuwe diagnostische technieken bij de ziekte van Alzheimer. *Ned Tijdschr Geneeskde* 1992; 136: 2459-62.
 38. Amaducci L, Angst J, Bech P, Benkert O, Bruinvels J, Engel RR. Consensus conference on the methodology of clinical trials of 'nootropics', Munich, June 1989. Report of the Consensus Committee. *Pharmacopsychiatry* 1990; 23: 171-5.
 39. CPMP. *Note for guidance: antidementia medicinal products*. CPMP Working Party on Efficacy of Medicinal Products. III/3705-91-EN, Draft 5. Brussels: Commission of the European Communities, 1992.
 40. Menges K. Proof of efficacy of nootropics for the indication "Dementia" (Phase III) - Recommendations. *Pharmacopsychiatry* 1992; 25:126-35.
 41. Oswald WD, Fleischmann UM. *Nürnberger Alters-Inventar (NAI)*. Manual. Nürnberg: Universität Erlangen-Nürnberg, 1993.
 42. Reitan RM. The relation of the Trail Making Test to organic brain damage. *J Cons Psychol* 1955; 19: 393-4.
 43. Van der Vlugt H. *Wechsler Memory Scale-Revised (WMS-R)*. Nederlandse bewerking (preliminary edition). Tilburg: Katholieke Universiteit Brabant, 1993.

44. Wechsler D, Stone CP. *Manual: Wechsler Memory Scale*. New York: Psychological Corporation, 1983.
45. Kalverboer AF, Deelman BG. *Vijftien Woordentest (15 WT)*. Nederlandse bewerking. Groningen: Rijksuniversiteit Groningen, 1984.
46. Shader RI, Harmatz JS, Salzman C. A new scale for clinical assessment in geriatric populations: Sandoz Clinical Assessment-Geriatric (SCAG). *J Am Geriatr Soc* 1974; 22: 107-13.
47. Venn RD. The Sandoz Clinical Assessment-Geriatric (SCAG) scale. A general-purpose psychogeriatric rating scale. *Gerontology* 1983; 29: 185-98.
48. Zieseniß E. *Auditbericht der Studie The Maastricht Ginkgo Trial*. Dortmund: GQS, Gesellschaft für Qualitätssicherung in der klinischen Forschung mbH, 1995.
49. Kieser M, Reitmeir P, Wassmer G. Test procedures for clinical trials with multiple endpoints. In: Vollmar J, ed. *Testing principles in clinical and preclinical trials*. Stuttgart: Gustav Fischer Verlag, 1995.
50. SPSS for Windows [program]. Version 7.5. Chicago: SPSS Inc., 1996.
51. Moran MA, Engelman L, Fitzgerald G, Lynch B. Polychotomous stepwise logistic regression. In: Dixon WJ, ed. *BMDP Statistical Software Manual*. Berkeley: University of California Press, 1990.
52. Arcus Quickstat. Biomedical version 1.0 (beta-version; build 82). [Online] <URL><http://www.camcode.com> [program]. 1.0 version. Cambridge: Research Solutions, 1997.

Appendix 1. Co-medication and exclusion

Persons using the following medications (Dutch trade names) were excluded from the trial:

- Antiparkinsonism medication: dopaminergic substances, e.g. levodopa (L-dopa in Madopar[®], Sine-met[®]), bromocryptine (Parlodol[®]), amantadine (Symmetrel[®]); anticholinergic substances, e.g. orfe-nadrine (Disipal[®]), biperideen (Akineton[®]), benztropine (Cogentin[®]), trihexyfenidyl (Artane[®])
- Antipsychotics (psychotropic substances): e.g. alimemazine (Nedeltran[®]), levomepromazine (Nozi-nan[®]), promazine ((Prazine[®]), thioridazine (Melleril[®]), chloorpromazine (Largactil[®]), periciazine (Neuleptil[®]), haloperidol (Haldol[®]), trifluoperazine (Trilafon[®]), pipamperon (Dipiperon[®]), pimozide (Orap[®]), zuclopentixol (Cisordinol[®])
However, with respect to some frequently used preparations (e.g. Haldol[®], Melleril[®], Nedeltran[®]) a decision will be made at the individual level.
- Neuroleptics: e.g. mestinon
- Reserpin treatment
- Cholinergic therapy: administration of metabolic precursors of acetylcholine (e.g.: choline, lecithine), cholinesterase-inhibitors (e.g. physostigmine, tacrine, HP029), or receptor-antagonists (e.g.: RS-86, betanechol, AF102B)
- Catecholamins: clonidine, zimeldine, mCPP, levodopa, selegiline and other substances with a nora-drenergic, serotonergic or dopaminergic activity
- Neuropeptids: e.g. somatostatin, vasopressin, ORG 2766, naltrexon
- GABA-agonists: e.g.: THIP, beta-carbolins
- Nootropics, used for stimulating the neuronal metabolism: e.g. co-dergocrine (Hydergine[®]), pi-racetam, oxiracetam, pramiracetam, sabeluzole, flunarizin, naftidrofuryl, pyritinol, nooptil, other Ginkgo biloba-preparations
- Experimental causal therapy, e.g. with N-methyl-D-aspartate-antagonists, calciumantagonists, anti-oxidants (selegelin, acetyl-levocarnitin, lazaroids, etc), or growth-factors (e.g. NGF, ganglioside GM1)
- Abuse of alcohol, drugs or pharmacologic substances

Persons using the following medications were included conditionally (if they were in a 'steady state' condition, if medication had not been started recently):

- Antidepressive therapy, e.g. clomipramine (Anafranil[®]), amitriptyline (Tryptizol[®]), imipramine (Tofra-nil[®]), doxepine (Sinequan[®])
- Antihypertensive therapy, e.g. methyl-dopa (Aldomet[®], Sembrina[®]), clonidine (Catapresan[®]); beta-blocking agents, such as propanolol (Inderal[®]), pindolol (Viskeen[®]), metoprolol (Selokeen[®], Lopres-sor[®]), atenolol (Tenormin[®]); diuretics
- Antiarrhythmics: e.g. digitalis-preparations (Lanicor[®], Lanoxin[®])
- Vaso-active, rheological active substances: e.g. Ascal-powder
- Antirheumatics (NSAID's): e.g. indometacine (Indocid[®]), ibuprofen (Advil[®], Brufen[®]), naproxen (Naprosyne[®])
- H₂-antagonists: e.g. cimetidin (Tagamet[®])
- Corticosteroids: e.g. Prednison, Oradexon[®], Kenacort[®]
- Analgetics, e.g. morfine (MS Continu[®]), bezitramide (Burgodin[®]), pentazocine (Fortral[®])
- Anti-epileptics: e.g. fenytoïne (Diphantoïne[®]), carbamazepine (Tegretol[®])
- Hypnotics, sedativa, tranquillizers (benzodiazepins): e.g. flurazepam (Dalmadorm[®]), nitrazepam (Mogadon[®]), lormetazepam (Noctamid[®]), temazepam (Normison[®]), flunitrazepam (Rohypnol[®]), di-azepam (Valium[®], Stesolid[®]), chloordiazepoxide (Librium[®]), lorazepam (Temesta[®]), medazepam (Nobrium[®]), oxazepam (Seresta[®])

These medications were allowed, provided that their use was limited to the night-time and that they had a fast turn-over rate, and therefore were unlikely to produce severe 'hang-overs' (e.g. triazolam, oxaze-pam, lorazepam, lormetazepam).

Appendix 2. Co-morbidity and exclusion

Diseases and conditions that led to exclusion were:

- Various neurological disorders, such as M. Parkinson, multiple sclerosis, chorea Huntingtoni, amyotrophic lateral sclerosis, vertebralis basilaris insufficiency, M. Pick, Jacob-Creutzfeld disease, M. Binswanger, epilepsy, myodystrophia, myasthenia gravis.
CVA's and TIA's will be reason for exclusion only when they have occurred recently (< 3 months ago) and depending on the remnant phenomena
- Traumata: normal pressure hydrocephalus (NPH), commotio cerebri, contusio cerebri, subdural hematoma
- Intoxication caused by heavy metal exposure, CO-exposure, alcohol abuse
- Tumours: intracranial primary brain tumours (malign or benign), ventricular cysts, intracranial haemorrhage, paracarcinomatous encephalopathy with psychical consequences (caused by hypoxia, edema, or toxic substances; e.g. associated with lung cancer, M. Hodgkin, M. Whipple, leukemia
- Severe infectious diseases and inflammations: intracranial, e.g., meningitis, encephalitis, brain abscess; or extracranial, e.g. AIDS, endocarditis lenta, lues, tuberculosis.
The presence of other types of infection will be documented and will give rise to exclusion only when indicated (e.g. urinary tract infections, respiratory tract infections)
- Severe vitamin deficiencies: vitamin B₁₂ (pernicious anemia), thiamin (Wernicke-Korsakov syndrome), folate
- Severe organic diseases: renal insufficiency, liver dysfunction, heart/circulation insufficiency (uncontrolled arrhythmias, cardiac failure causing hypoxia, anoxia, ischemia), pulmonary insufficiency
- Severe gastrointestinal disorders, which might prevent the test substance from being absorbed
- Aphasic, sensoric, motoric and visual disorders which might interfere with the measurement procedures
- State of severe depression
- Psychotic disease: transient organic psychosis (e.g. delirium), schizophrenic type of psychosis, affective psychosis (e.g. depressive pseudodementia)
- State of impaired consciousness

Medical conditions that required exclusion in some cases, were:

- Certain metabolic disorders, e.g., hyperthyroidism, hypothyroidism, hyperparathyroidism, M. Cushing, electrolyte disturbances (dehydration, hyperhydration, K/Na-imbalance, hypercalcaemia, Wilson's disease)
People suffering from diabetes mellitus are allowed to participate, provided that their condition has been stabilized and controlled for
- Conditions causing hypoxia, anoxia or ischemia: severe anemia, severe hypotension or hypertension, etc.

People with carotis stenosis were only excluded if surgery was indicated.

Appendix 3. Overview of the selection procedure

Participants of the trial were selected in several steps. The table on the next page gives detailed information on the numbers of residents in the participating homes for the elderly who were included as candidates in the various steps of the selection procedure.

The source population consisted of 4,459 residents in 46 homes for the elderly. Actually 39 homes for the elderly cooperated, but 7 of them participated twice. The overview shows per step in the procedure the numbers of candidates who 'survived' a step and how many were excluded for the next step, and for what reason. The columns in the overview refer to the following steps in the procedure:

- How many residents were initially preselected by the staff of the homes?
- How many of them fulfilled the (preliminary) inclusion and exclusion criteria and could thus be contacted by the staff for participation?
- How many subjects were screened by the researchers?
- How many residents still fulfilled the selection criteria after the screening and could be included in the run-in-period?
- How many residents were included in the baseline measurements?
- Finally, how many subjects were randomised and definitively included in the trial?

The reasons for exclusion are described shortnoted. For instance, "166 refused" in the column "selected after check criteria" indicates that 166 residents decided not to participate in the screening interview. The criteria regarding the Syndrom Kurz Test (SKT, measured during the screening interview) were modified during the study. Up to home 8 a SKT score of 9-18 was needed for subjects with the 'diagnosis' AAMI, whereas those 'diagnosed' as having dementia needed a SKT score of 9-23. From home 9 onwards, a SKT score between 8 and 23 was needed, irrespective of an AAMI or dementia diagnosis. Subjective memory complaints were measured using the MAC-Q questionnaire, with scores ranging from 5 to 25 points. A score of at least 12 points was needed from either the resident or the staff of the home to fulfill the inclusion criterium regarding these complaints.

The last column of the table shows the distribution of the participants over the 4 strata that we used for the randomisation. These strata correspond with:

- Stratum 1: diagnosis AAMI and memory problems constant or improved over the last six months
- Stratum 2: diagnosis AAMI and memory problems deteriorated over the last six months
- Stratum 3: diagnosis dementia and memory problems constant or improved over the last six months

Stratum 4: diagnosis dementia and memory problems deteriorated over the last six months.

The Maastricht Ginkgo Trial was prepared during 1993. Recruitment of the participants started in January 1994. The first eligible subjects were randomized in March 1994. Intake of new study subjects continued until the beginning of June 1996. The last 24 week follow-up measurements were performed in November 1996.

Overview of the selection procedure for the Maastricht Ginkgo Trial

Source population (46 homes for the elderly)	Preliminary selected by staff	Selected after check criteria	Screened by researchers	Run-in-period	Baseline	Randomised and stratum
n = 4,459 (100%) (5%)	n = 781 (16%)	n = 705 (16%)	n = 513 (12%)	n = 256 (6%)	n = 215	n = 214*
Reason for exclusion per step in procedure:	<ul style="list-style-type: none"> 40 co-medication 21 co-morbidity 1 visus too bad 6 excluded in co-operation with GP 3 not clear 3 deceased 1 alcohol abuse 1 nursing home admission 	<ul style="list-style-type: none"> 166 refused 6 family refused 3 co-medication 1 deceased 1 recent psychose 1 deterioration health 2 GP refused 2 longterm hospital admission 4 'poor medical condition' 1 excluded (probably non-complier) 5 other reasons 	<ul style="list-style-type: none"> 90 SKT <8/9 27 SKT >23 8 SKT >18 and non-dementia 22 MAC-Q <12 3 co-medication 29 interview impossible 10 drop-out during interview 34 refused after all 5 co-morbidity 4 GP refused after all 10 visus too bad 4 'poor medical condition' 2 hearing deficiency 1 severe cognitive disorder (according to staff) 4 depression 1 psychiatric disorder 1 hospital admission 2 too confused 	<ul style="list-style-type: none"> 1 partner deceased 4 refused medication after all 3 recent co-medication 1 poor prognosis 2 longt. hospital admission 2 too confused 2 family refused after all 2 non-complier 7 co-morbidity 1 problems with medication intake 1 problems with partner 13 refused after all 1 placebo responder 1 GP refused after all 	<ul style="list-style-type: none"> 1: n=89 2: n=62 3: n=41 4: n=22 	
Total number of excluded residents per step	76 residents	192 residents	257 residents	41 residents		

* All participants included in the baseline measurements were randomised. Due to an administrative mistake, one person was randomised erroneously. As a consequence, the final study population consisted of 214 participants.

4

Ginkgo for elderly people with dementia and age-associated memory impairment: results of a randomized controlled trial

Martien CJM van Dongen, MSc¹

Erik van Rossum, PhD¹

Alphons GH Kessels, MD¹

Hilde JG Sielhorst, RN^{1, 2}

Paul G Knipschild, MD, PhD²

¹ Department of Epidemiology, Maastricht University

² Department of General Practice, Maastricht University

ABSTRACT

Background

Many studies have suggested the clinical efficacy of Ginkgo extract as a nootropic and anti-dementia drug. We (re-)investigated its efficacy, the dose-effect relationship, and the effect durability in a randomized clinical trial.

Methods

In 39 homes for the elderly in the Netherlands, we randomly assigned 214 elderly patients with dementia or age-associated memory impairment (AAMI) to three dosages of orally administered *Ginkgo biloba* special extract (EGb 761): 240 mg/day, 160 mg/day, and 0 mg/day (placebo). The intervention period was 24 weeks. After 12 weeks the subjects who received Ginkgo were randomly allocated once again, to either continued Ginkgo treatment or placebo treatment. Primary outcomes were psychometric functioning (Syndrom Kurz Test (SKT), scale of change in score: -27 to +27), psychopathology (clinical global impression (CGI-2), scale: 1 to 7), and behaviour (instrumental activities of daily life (NAI-NAA), scale of change in score: -40 to +40).

Findings

123 Patients received either Ginkgo (240 or 160 mg/day, n=79) or placebo (n=44) during the entire intervention period. We found no statistically significant differences in mean change of scores over this 24 week period, neither in an overall test ($T_{OLS} = 0.34$; $p=0.37$), nor for each of the outcome measures separately. The differences were: SKT: +0.4 (90% CI: -0.9 ; +1.7); CGI-2: +0.1 (90% CI: -0.3 ; +0.4), and NAI-NAA: -0.4 (90% CI: -1.9 ; +1.2). A positive difference is in favour of Ginkgo. Neither the dementia sub-group (n=36), nor the AAMI subgroup (n=87) experienced a significant effect of Ginkgo treatment. There was no dose-effect relationship, and neither was there an effect of prolonged Ginkgo treatment. We also did not find an association between the study treatment and the occurrence of (serious) adverse events.

Interpretation

The results of this trial do not support the view that *Ginkgo biloba* is beneficial for patients with dementia or age-associated memory impairment.

1 INTRODUCTION

Efforts to identify substances which stand out for both their efficacy and safety in treating dementia and related cognitive disorders, have yielded disappointing results until now. Several years ago the *Lancet* published a criteria-based literature review reporting a rather consistent beneficial effect of Ginkgo for patients with cerebral insufficiency¹. Forty randomized controlled trials were included in that review, 39 of which showed a statistically significant effect or at least a positive trend for Ginkgo. However, the methodological quality of many trials was considered to be poor. Moreover, the studies entailed a very heterogeneous collection of target health problems, ranging from overt dementia to non-cognitive manifestations of brain dysfunction, such as vertigo and tinnitus.

Medicinal Ginkgo preparations are obtained through extraction of dried leaves of the *Ginkgo biloba* tree. *Ginkgo biloba* special extracts, such as EGb 761 and LI 1370, have been subjected to a standardized process of enrichment and purification, yielding fixed concentrations of two active substance classes (Ginkgo flavonol glycosides: 16-25%; terpene trilactones: usually 6%). These extracts have been registered as drugs for various indications in several European countries. Ginkgo is marketed in these countries under various brand names, and has become a commercial best seller. Ginkgo monopreparations and multipreparations with Ginkgo constituents have also attracted a growing popularity as over-the-counter drugs^{2,3}. Several mechanisms of action have been described to explain the claimed nootropic properties of Ginkgo: increased tolerance to hypoxia; prevention of membrane damage through free radical scavenger activity; improvement of blood rheology and vasoregulating capacity, resulting in an increased blood flow; preven-

tion of post-traumatic or toxin-induced brain edema; platelet activating factor inhibition and neuroprotective action, either by direct or by indirect - modulated through blood flow - influences on the nervous system⁴.

We initiated a trial to retest the hypothesis that Ginkgo is beneficial to elderly people with dementia or age-associated memory impairment. The promising results of previous trials justified the inclusion of two additional research questions, one dealing with the dose-dependency and the other with the durability of any Ginkgo effect.

2 METHODS

2.1 STUDY DESIGN

The trial was designed as a multi-centre, randomized, placebo-controlled, double-blind trial (see figure 1: trial profile). The eligible subjects were randomly allocated to one of three treatments: EGb 761 160 mg/day ('usual dose'), 240 mg/day ('high dose'), or 0 mg/day (placebo). This allowed us to study the dose-effect relationship as well. The participants who completed the first 12-week treatment period under Ginkgo, were randomly allocated again for a second 12-week treatment period, either to continue their Ginkgo treatment or to receive placebo. Initial placebo use was continued during the second intervention period. This allocation scheme for the second intervention period allowed us to study the durability of the intervention effect after discontinuation of Ginkgo therapy. For the randomization procedure we used a computer generated random sampling set. Random allocation was done within four predefined strata, according to diagnostic subcategory (dementia vs AAMI) and perceived memory change during the 6-month pretreatment phase (improvement or no change vs impairment). The first random allocation procedure (week 0) was con-

ducted in blocks of five patients and with a 2 : 2 : 1 ratio (Ginkgo 240 mg : Ginkgo 160 mg : placebo). By the end of the study this was changed to 1 : 1 : 1, to add statistical power for analysing the primary research question. The second randomization procedure (week 12) used blocks of 2 patients and a 1 : 1 (Ginkgo : placebo) allocation ratio.

The entire 24-week treatment period was preceded by a 3-week placebo run-in period. Eventually, the repeated random allocation procedure yielded five different treatment groups, of approximately equal size: 1. Ginkgo 240 mg (24 weeks); 2. Ginkgo 160 mg (24 weeks); 3. placebo (24 weeks); 4. Ginkgo 240 mg (12 weeks) + placebo (12 weeks); 5. Ginkgo 160 mg (12 weeks) + placebo (12 weeks). This paper focuses on the first three groups.

For a more detailed description of the design, the conduct, and the results of the Maastricht Ginkgo Trial we refer to the official trial report⁵.

2.2 STUDY POPULATION

The target population consisted of patients from two diagnostic categories. Firstly, we included subjects with a diagnosis of uncomplicated dementia according to the DSM-III-R⁶ and ICD-10 criteria⁷. Patients with either Alzheimer dementia (AD), vascular dementia (VD), or a mixed type of dementia were eligible. Secondly, we enrolled subjects suffering from age-associated memory impairment (AAMI), according to a slightly adapted version of the diagnostic criteria proposed by Crook et al⁸. We decided to include both AAMI patients and patients with beginning dementia in our study, since we assumed that relatively mild stages of cognitive decline will provide for the clearest manifestations of any Ginkgo effect, and since there are arguments to consider AAMI a precursor lesion of full-blown dementia.

From January 1, 1994 to May 31, 1996 we recruited study participants from 39 old people's homes in the southern part of the Netherlands, applying a multi-stage screening procedure. For the differential diagnosis of dementia we used the SIDAM⁹. This is a Germany-based semi-structured interview, which includes the Mini-Mental State Examination (MMSE)¹⁰ and the Hachinski Ischemic Score (HIS)¹¹. It keeps close track of the DSM-III-R criteria for dementia. The degree of severity of dementia had to be mild to moderate as assessed by the SIDAM interview and by a score between 8 and 23 on the Syndrom Kurz (Syndrome Short) Test (SKT), a psychometric test for the assessment of memory and attention¹². For the AAMI diagnosis subjects had to suffer from impaired cognitive functioning, both objectively and subjectively, in the absence of dementia. This was indicated by a score between 8 and 23 on the SKT, combined with the presence of subjective memory complaints (total score ≥ 12 assigned by either the study subject or the nursing staff, using a modified version of the Memory Assessment Clinics Questionnaire (MAC-Q)¹³, a 5-item memory questionnaire which refers to important memory problems in daily life).

In addition subjects of both diagnostic subgroups had to fulfil the following enrolment criteria: age ≥ 50 years; written informed consent by patient and/or legal representative; absence of severe depression (Geriatric Depression Scale (GDS), 15-item version, total score < 11)^{14,15}; presence of an adequate level of premorbid intelligence (IQ > 80 , global assessment); sufficient compliance; absence of placebo-response during the run-in period; no expectation of premature withdrawal; absence of serious co-morbidity, in particular of pathological conditions that are considered nontreatable underlying causes of dementia; absence of sources of interference with the trial conduct (e.g., various neurological disorders, brain traumata,

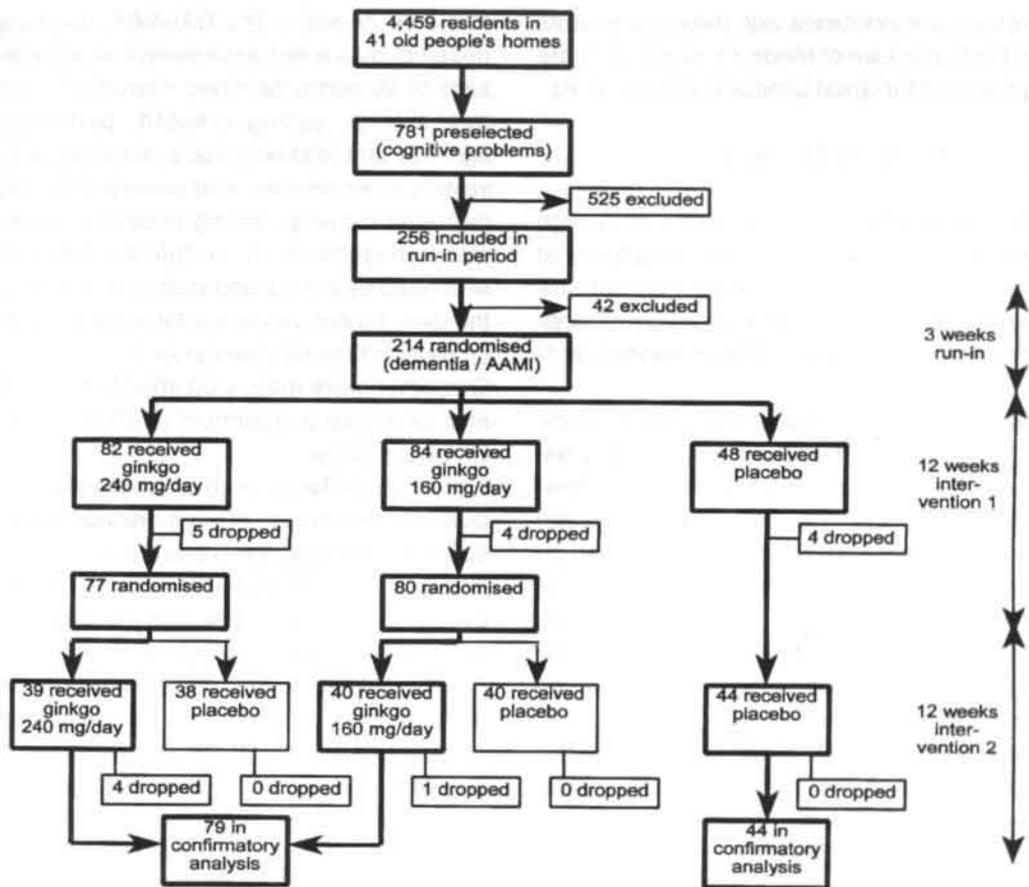


Figure 1. Trial profile.

tumours, severe infectious diseases; absorption disorders); absence of impermissible co-interventions, in particular drugs with a debilitating influence on cognitive functioning or with a claimed nootropic action (e.g., antipsychotic drugs, antiparkinson medication, neuroleptics, cholinergic therapy, antidepressants, vasoactive drugs).

2.3 STUDY TREATMENTS

Ginkgo was administered orally through two film-coated tablets per day, each containing either 80 mg or 120 mg of the extract, standardized on 24% Ginkgo flavonol glycosides

and 6% terpene trilactones (Ginkgolides, bilobalid). The placebo consisted also of 2 film-coated tablets per day, but these contained no Ginkgo extract. We made special efforts to check the comparability of the verum and the placebo tablets with respect to appearance, colour, smell, taste, granularity and solubility. In order to mimic the typical taste of Ginkgo, 2 mg of quinine hydrochloride was added to the placebo tablets. The study participants consumed their daily dose at fixed times (morning and evening, during the meal), mostly under direct supervision of a nursing staff member. Packaging and labelling of the study medication were per-

formed in accordance with the requirements of the Dutch Law of Medicine Supply and the guidelines for Good Clinical Practice (GCP).

2.4 STUDY OUTCOMES

To evaluate the intended effects of Ginkgo we defined primary outcome measures at three different levels, in agreement with recent standards for the evaluation of anti-dementia drugs and cognition enhancers¹⁶⁻¹⁸.

Cognitive functioning at the psychometric level (neuropsychological testing) was assessed by the SKT. This test battery, focussing on memory (perceptual motor speed) and attention, was conducted by trained interviewers. The SKT consists of 9 one-minute subtests which are partly speed-oriented and partly span-oriented (e.g., naming objects, arranging blocks, counting symbols, delayed recall of objects, recognition of objects). In our trial we used three parallel versions. Scaled subtest scores are aggregated to a SKT total status score ranging from 0 ('very good') to 27 ('very poor').

Change in clinical signs (psychopathology) over time was assessed by the Clinical Global Impression, item 2 (CGI-2). This rating instrument expresses the global change in observable cognitive functioning directly on a transitional scale ranging from 1 ('very much improved') to 7 ('very much deteriorated'). The CGI-2 was rated by the nursing staff in each old people's home.

We measured cognition-related behaviour by the Nürnberger Alters Alltagsaktivitäten Skala (NAA; Nuremberg Gerontopsychological Rating Scale for Activities of Daily Living), which is part of the Nürnberger Alters Inventar (NAI; Nuremberg Gerontopsychological Inventory). The NAI is a German battery of psychometric tests, questionnaires and observational rating scales, aimed at the measurement of cognitive functioning, performance of daily life activities, well-being

and care needs¹⁹. The NAI-NAA, originally developed as a self-assessment scale, consists of 20 items, covering instrumental activities (e.g., getting dressed, performing light household tasks), social activities (e.g., visiting other people), and general cognitive performance (e.g., writing postcards, financial management). In our trial the NAI-NAA was rated by the nursing staff. Summation of the item scores yields a total score from 20 ('very good') to 60 ('very poor').

Outcomes were measured after 4, 8, 12, 18, and 24 weeks of treatment (NAA at 12 and 24 weeks only).

The study protocol listed also several other outcome measures. We will discuss the results for these measures elsewhere. We registered the occurrence of (serious) adverse events periodically. For this purpose we used among other things a checklist of health complaints. This was filled out by the nursing staff at each measurement time point. We also monitored the exposure to co-interventions and the compliance with the study medication (pill counting).

2.5 STATISTICAL ANALYSIS

The confirmatory analysis of the study results was an intention-to-treat analysis comparing the 0-24 week changes in the scores on the three primary outcome measures between the subgroups who received Ginkgo (160 mg and 240 mg combined) and placebo during the entire 24-week treatment period. The hypothesis testing system was headed by a global multivariate OLS-test²⁰ of the null-hypothesis that the subjects treated with EGb 761 and placebo would not differ for either of the three primary outcome measures. This overall test was followed by separate tests for each of the outcome measures, in a hierarchical order (first SKT, next CGI-2, and finally NAA).

We also performed various descriptive analyses, multivariate analyses to adjust for

baseline differences between the comparison groups, comparisons between the high-dose and the normal-dose Ginkgo groups, comparisons between groups with prolonged versus interrupted Ginkgo use, 'per protocol' analyses restricted to the 'valid' cases only, and various subgroup analyses. In addition to the overall OLS-test ($\alpha = 0.05$; one-sided), the statistical procedures included unpaired t-tests ($\alpha = 0.05$; one-sided) of the changes in SKT and NAA (continuous), Mann-Whitney U-tests of the differences in the CGI (categorical), multiple linear regression analyses to estimate the adjusted effect of study treatment on change in SKT and NAA (controlling for baseline differences in scores on several variables: age, gender, diagnostic subcategory (AAMI vs dementia), pretreatment level of change in memory complaints, baseline value of the outcome measure concerned, assessed quality of old people's home, level of co-interventions during follow-up), and polytomous logistic regression analyses to estimate the adjusted effect on CGI (controlling for the same set of independent variables). All participants with baseline and 24-week follow-up scores on the primary outcome variables were included in the intention-to-treat analysis. Missing values were substituted according to a predefined replacement scheme, with the last-observation-carried-forward principle as the main feature. To qualify for the per protocol analysis study subjects had to fulfil minimum requirements regarding compliance ($\geq 75\%$ intake of prescribed medication), adherence to the outcome measurement time table, and (non-)exposure to unpermitted co-interventions. Moreover, we performed a responder analysis. Study subjects with a clinically meaningful response to the study treatment for at least two of the three primary outcome variables were regarded responders. Changes $\geq +3$ for the SKT, and $\geq +4$ scale-points for the NAA, and a score < 4 for the CGI-2 after 24

weeks of follow-up, were considered clinically meaningful improvements. Subgroup analyses and a comparative analysis of the number and type of (serious) adverse events completed the data analysis.

3 RESULTS

39 Old people's homes participated in the trial, 7 of them twice. They harboured 4,459 residents, 781 (18%) of whom were preselected by the nursing staff for presumed cognitive problems. We excluded subsequently 268 persons at an early stage of the patient selection procedure, mainly because of co-morbidity, co-medication and refusal by the patient, while 513 (12%) proceeded to the formal screening interview. Of these 256 (6%) qualified for the run-in period, and finally 214 patients met the remaining enrolment criteria and agreed to sign informed consent. Ineligible scores on the screening tests (SKT, MAC-Q), problems with completing the interview, a poor visus or hearing function, several other types of co-morbidity, and refusals were the major reasons for not entering the run-in period. By the end of the run-in period some 20 candidates refused to participate after all, 11 suffered from serious co-morbidity or a poor prognosis, 3 appeared to be non-compliers, 1 a placebo-responder, 2 were admitted to hospital, and 3 used banned co-medication.

The eligible subjects, 63 with dementia and 151 with AAMI, were randomly allocated to the treatment regimens.

Thirteen of the study participants dropped out during the first 12 weeks, leaving 201 patients to be randomized for the second 12-week intervention period. Another 5 subjects withdrew during the second intervention phase. The dropouts were distributed as follows: Ginkgo 240 mg: 9/82 (11%); Ginkgo 160 mg: 5/84 (6%); placebo: 4/48 (8%). The reasons for attrition were: death (8x), hospi-

Table 1. Distribution of baseline characteristics, by intervention subgroup (placebo 0-24 weeks (n=44) vs Ginkgo 160/240 mg 0-24 weeks (n=79)). Absolute numbers (percentages) for categorical variables, and mean values (standard deviations) for (semi-)continuous variables.

		placebo (n=44)	Ginkgo (n=79)
Gender	male	8 (18%)	11 (14%)
	female	36 (82%)	68 (86%)
Communal living situation	alone	39 (89%)	70 (89%)
	together	5 (11%)	9 (11%)
Level of cognitive impairment	mild	20 (46%)	33 (42%)
	moderate	24 (55%)	46 (58%)
Perceived memory complaints (6 month pretreatment period)	increased	16 (36%)	32 (41%)
	the same / decreased	28 (64%)	46 (59%)
State of dementia	dementia (mild/moderate)	13 (30%)	23 (29%)
	AAMI	31 (71%)	56 (71%)
Age (years)		82.5 (5.8)	82.6 (5.1)
Quetelet Index (kg/m ²)		24.9 (3.7)	24.9 (3.3)
MMSE (Mini-Mental State)		18.7 (4.6)	18.0 (4.9)
MAC-Q, modified, assessed by patient (subjective memory complaints)		12.7 (2.8)	12.2 (3.0)
MAC-Q, modified, assessed by nursing staff (subjective memory complaints)		15.2 (4.6)	16.4 (4.3)
SKT (memory and attention)		14.1 (4.6)	15.6 (4.1)
NAI-NAA, assessed by nursing staff (activities of daily life)		42.0 (8.5)	44.3 (7.2)

Scale ranges (underlining indicates the most favourable extreme); MMSE: 0-30; MAC-Q: 5-25; SKT: 0-27; NAI-NAA: 20-60

tal admission or caregiver request (3x), serious concurrent illness (4x), deliberate avoidance of further cognitive testing (3x).

Of the 214 randomized study subjects, 123 qualified for the confirmatory analysis, which compared Ginkgo (240, 160 mg/day) with placebo for 24 weeks. Table 1 demonstrates no systematic baseline differences between the combined Ginkgo groups (n=79) and the placebo group (n=44).

Unadjusted effects of Ginkgo treatment are summarized in table 2. Both the Ginkgo group and the placebo group showed a minor deterioration over the 24 week follow-up period. The differences between groups

were small to non-existent, and not statistically significant for either of the outcome measures (SKT: +0.4; CGI-2: +0.0; NAI-NAA: -0.0).

The result of the multivariate global OLS-test does not allow us to reject the null-hypothesis: $T_{OLS} = 0.34$ ($p = 0.37$). In this test T_{OLS} is the sum of the three T-statistics shown in table 2, divided by the square root of the sum of the Pearson correlation coefficients for each pair of primary outcome measures.

Multiple linear regression analysis to control for potential confounding caused by baseline differences, yielded similar estimates of the

Table 2. Comparison of the mean baseline scores, mean 24-week follow-up scores, and mean week 0-24 changes in score for SKT, CGI-2 and NAA, for the placebo group (n=44), the Ginkgo high dosage group (n=39), and the Ginkgo usual dosage group (n=40).

Outcome variable	Intervention subgroup	Mean 0 w. baseline score †	Mean 24 w. follow-up score †	0-24 w. change in score ††: mean (SD) *	0-24 w. change **: mean (90% CI)	T-statistic (P-value, one-sided)
Memory and attention (SKT)	placebo	14.1	15.3	-1.2 (3.8)		0.53
	Ginkgo 160 + 240	15.6	16.4	-0.8 (4.1)	0.4 (-0.9; +1.7)	(0.30)
	Ginkgo 160	15.8	16.4	-0.7 (4.4)		-0.32
	Ginkgo 240	15.5	16.5	-1.0 (3.9)	-0.3 (-1.9; +1.3)	(0.37)
Clinical global impression (CGI-2)	placebo		4.3	(1.1)		0.23
	Ginkgo 160 + 240		4.3	(1.2)	0.0 (-0.3; +0.4)	(0.41)
	Ginkgo 160		4.2	(1.1)		1.10
	Ginkgo 240		4.4	(1.2)	-0.3 (-0.7; +0.2)	(0.14)
Activities of daily life (NAA)	placebo	42.0	43.3	-1.4 (5.5)		-0.00
	Ginkgo 160 + 240	44.3	45.7	-1.4 (4.8)	-0.0 (-1.6; +1.6)	(0.50)
	Ginkgo 160	44.0	45.2	-1.2 (4.9)		-0.27
	Ginkgo 240	44.7	46.2	-1.5 (4.8)	-0.3 (-2.1; +1.5)	(0.40)

* For CGI-2: SD 24 w. follow-up score

** For CGI-2: 24 w. follow-up difference: mean (90% CI)

† Range of status scores (underlining indicates the most favourable extreme): SKT: 0-27; CGI-2: 1-7; NAA: 20-60†† Range of change scores (underlining indicates the most favourable extreme): SKT: -27 - +27; NAA: -40 - +40

effects regarding SKT and NAA. For SKT the adjusted difference in change over 24 weeks was +0.4 scale points (90% CI: -0.9; +1.7), in favour of Ginkgo, which is equal to the unadjusted effect. Adjustment for confounding resulted in a shift of the NAA-difference to -0.4 (90% CI: -1.9; +1.2), in favour of placebo. Polychotomous regression analysis on the 24 week CGI-2 score yielded an estimated odds ratio of 1.4 (95% CI: 0.6; 3.0), indicating that subjects in the Ginkgo group had 1.4 times more chance to have a favourable shift in their cognitive functioning recorded by the nursing staff. Again, this increase in chance was not statistically significant.

The descriptive analysis of the 12-week intervention results revealed for SKT an effect of +1.0 (90% CI: +0.0; +2.0), for CGI of -0.0 (90% CI: -0.3; +0.2), and for NAA of +0.4

(90% CI: -0.8; +1.7). T-testing produced a (borderline) significant effect for SKT only ($t=1.69$; $p=0.05$). Multiple linear regression analysis yielded regression coefficients corresponding with adjusted 12-week treatment effects of +1.0 (90% CI: -0.1; +2.1) and +0.3 (90% CI: -1.0; +1.5) for SKT and NAA, respectively. Polychotomous logistic regression analysis on the 12-week CGI-2 score yielded an estimated OR of 1.3 (95% CI: 0.6; 2.7).

Table 2 also illustrates that a dose-effect relationship is lacking, which is not surprising, given the absence of a beneficial effect for the combined Ginkgo groups. The results of neither a non-parametric Kruskal-Wallis test nor more powerful tests for trend supported a dose-related effect of Ginkgo. An equivalent analysis of the 12-week follow-up data yielded similar findings, except for a signifi-

Table 3. Comparison of the mean baseline scores, mean 24-week follow-up scores, and mean week 0-24 changes in score for SKT, CGI-2 and NAA, Ginkgo 160/240 + Ginkgo 160/240 (n=79) vs Ginkgo 160/240 + placebo (n=78) vs placebo + placebo (n=44).

Variable	Intervention subgroup	Mean 0 w. baseline score †	Mean 24 w. follow-up score †	0-24 w. change in score ††: mean (SD)*	F-statistic	0-24 w. change** difference ††: mean (90% CI)	T -statistic (P-value, one-sided)
Memory and attention (SKT)	placebo + placebo	14.0	15.3	-1.2 (3.8)	0.17	0.2 (-0.8; +1.2)	0.38 (0.35)
	Ginkgo + placebo	15.1	16.1	-1.1 (3.4)			
	Ginkgo + Ginkgo	15.6	16.4	-0.8 (4.1)			
Clinical global impression (CGI-2)	placebo + placebo		4.3	(1.1)	0.10	0.1 (-0.2; +0.4)	-0.45 (0.33)
	Ginkgo + placebo		4.4	(1.1)			
	Ginkgo + Ginkgo		4.3	(1.2)			
Activities of daily life (NAA)	placebo + placebo	42.0	43.3	-1.4 (5.5)	0.25	-0.5 (-1.7; +0.7)	-0.67 (0.25)
	Ginkgo + placebo	43.2	44.1	-0.9 (4.5)			
	Ginkgo + Ginkgo	44.3	45.7	-1.4 (4.8)			

* For CGI-2: SD 24 w. follow-up score

** For CGI-2: 24 w. follow-up difference: mean (90% CI)

† Range of status scores (underlining indicates the most favourable extreme): SKT: 0-27; CGI-2: 1-7; NAA: 20-60

†† Range of change scores (underlining indicates the most favourable extreme): SKT: -27 - +27; NAA: -40 - +40

Table 4. Number of patients with clinically meaningful improvements for SKT, CGI and NAA, after 24 and 12 weeks of follow-up (intention-to-treat population).

Number of outcome variables with clinically meaningful improvement after 24 weeks *	placebo + placebo		Ginkgo 160 + Ginkgo 160		Ginkgo 240 + Ginkgo 240		Ginkgo 160 + Placebo		Ginkgo 240 + Placebo	
	no.	%	no.	%	no.	%	no.	%	no.	%
	3 variables	1	2.3%	1	2.5%	1	2.6%	-	0.0%	2
2 variables	7	15.9%	6	15.0%	3	7.7%	2	5.0%	4	10.5%
1 variable	7	15.9%	14	35.0%	10	25.6%	15	37.5%	4	10.5%
0 variables	29	65.9%	19	47.5%	25	64.1%	23	57.5%	28	73.7%
Total	44	100.0%	40	100.0%	39	100.0%	40	100.0%	38	100.0%

* Clinically meaningful improvement: SKT: $\geq +3$ points; CGI-2: <4 points; NAI-NAA: $\geq +4$ points

cantly better result regarding the change in NAA-scores for the high-dose Ginkgo group compared with the low dose Ginkgo group. Table 3 shows that prolongation compared with discontinuation of Ginkgo use after 12 weeks is not accompanied by more beneficial 24-week outcomes.

As table 4 shows, 27 study participants could be classified as 'responder' after 24 weeks of treatment. The subgroup which received

Ginkgo during the entire period, harboured 14% responders (11/79), the placebo group 18% (8/44). A statistically significant association between responder status (yes / no) and intervention subcategory was absent ($\chi^2=0.39$; $p=0.53$). After 12 weeks of follow-up we counted 29 responders, viz. 14% (22/160) of the patients on Ginkgo, and 16% (7/44) on placebo.

The findings of the per-protocol analyses, involving 112 study subjects after 24 weeks and 187 after 12 weeks of treatment, did not alter the conclusions based on the intention-to-treat analyses.

Stratified analyses were conducted to identify subgroup-specific effects. Compared with demented patients, AAMI patients seemed to have benefitted slightly more from 24 weeks of treatment with either Ginkgo 240 mg or Ginkgo 160 mg. A difference in NAA-change, in favour of placebo, was found for the dementia subgroup: -3.4 (90% CI: -6.0 ; -0.9). However, this effect was not statistically significant anymore after adjustment for confounding: -1.9 (90% CI: -4.2 ; $+0.5$). A statistically significant effect on change in SKT-score, this time in favour of Ginkgo, was found for the AAMI subgroup, both before ($+1.3$; 90% CI: $+0.1$; $+2.6$) and after ($+1.4$; 90% CI: $+0.0$; $+2.7$) adjustment for confounding. Additional subgroup analyses were performed for age (<84 vs ≥ 84 years), MMSE baseline score (<19 vs ≥ 19 points), SKT baseline score ($8-15$ vs $16-27$ points), and the number of so-called biological life events (BLEs), such as history of general anaesthesia, history of brain damage, neurological disorders, and alcohol abuse, which are believed to have a negative impact on cognitive status ($0-2$ vs $2-5$ BLEs). We failed to identify a consistent effect of Ginkgo for any of these subgroups.

During the first intervention period 111 (52%) of the 214 study subjects were identified with at least one health complaint. The adverse event rates were 46/82 (56%), 44/84 (52%), and 21/48 (44%) for Ginkgo 240 mg, Ginkgo 160 mg, and placebo, respectively. Dizziness ($n=58$), nervousness ($n=49$), and headache ($n=32$) were the most frequently reported signs. During the second intervention period symptoms were reported for 113 (56%) of the 201 remaining study subjects, with the treatment specific rates being 18/39 (46%), 16/40 (40%), and 79/122 (65%), re-

spectively. A clear association with type of treatment did not exist. During the entire trial period 35 serious adverse events were registered: 8 deaths, 25 hospital admissions, and 2 cases who had developed a malignancy. These events occurred in 26/166 (16%) of the subjects initially treated with Ginkgo, and in 9/48 (19%) with placebo. For only one case - a placebo user - a possible association with study treatment was hypothesized.

We were surprised to find quite a high rate of symptoms that might be indicative for a withdrawal effect among the study subjects who shifted from Ginkgo to placebo after 12 weeks of intervention. After 18 weeks of follow-up at least one withdrawal symptom was registered for almost 50% of the ex-Ginkgo users but for only 25% of the subjects who still received Ginkgo. Unexpectedly, this two-fold difference was still present after 24 weeks of follow-up. The strongest effect was observed for three symptoms that are closely related to brain activity: fatigue, lack of concentration, and irritation. We failed to find a reasonable explanation for this phenomenon.

4 DISCUSSION

Our trial did not reveal any benefit of Ginkgo over placebo. With regard to SKT and CGI we found very small differences after 24 weeks of treatment, in favour of Ginkgo. These differences were neither statistically significant nor clinically meaningful. With respect to NAA the difference was also neglectable. Adjustment for possible confounding by several prognostic variables did not alter these results. After 12 weeks of follow-up a small, but statistically significant effect was registered for SKT, in favour of Ginkgo. The effects for NAA and CGI-2 were not statistically significant. Again, adjustment of the estimated effects by multivariate regression

analysis did not change the results. A responder analysis showed almost equal proportions of responders to Ginkgo and placebo, after both 24 and 12 weeks of intervention. We also did not find a dose-effect relationship or a persistence of any effect after discontinuation of Ginkgo treatment.

In order to assess to what extent various sources of bias may have produced spurious results we now like to reflect critically on the adequacy and comparability of the study populations, interventions, and outcome measures.

4.1 ADEQUACY

We have deliberately selected residents of old people's homes as the source population for our trial, in order to optimize the level of adherence to the study medication and the outcome measurement time schedule. This recruitment strategy has caused an overrepresentation of the very old (mean age: 84 years), and a high level of co-morbidity among the study participants. Moreover, we can imagine that the protective living conditions make residents of old people's homes feel less conscious and worried about their cognitive functioning. Apparently the provided care and shelter may extinguish some mental challenges which elderly people living in the 'open' community have to cope with. This may also have compromised the correct classification of eligible study subjects, since the assessment of the impact of cognitive abilities on the level of social functioning plays a decisive role in the differentiation between dementia and non-dementia forms of cognitive decline (see DSM-criterion C for dementia). By the end we believe that part of the study subjects who have been assigned to the AAMI-subgroup in our trial, actually would have deserved the diagnosis 'dementia'.

We have deliberately abstained from making a distinction between Alzheimer's dementia, vascular dementia, and other types of dementia. Given the non-clinical setting of our trial we had to face constraints as for the availability of medical information and the application of sophisticated diagnostic techniques. Moreover, until recently this kind of information was considered not very relevant for the prescription of Ginkgo - given its multiple mechanisms of action - and for the clinical management of dementia cases in continental Europe²¹.

We applied a high-quality, standardized Ginkgo extract (EGb 761), with a significant exposure contrast between verum treatment (either 2 x 80 or 2 x 120 mg of active substance per day) and placebo. Much effort was paid to develop a valid placebo preparation. The distribution schedule, supervision by the nursing staff, and periodical pill counts were major safeguards against an irregular or insufficient level of exposure to the active Ginkgo compounds. Potential non-compliers were filtered out already during the run-in phase. Unwanted interference with co-interventions was precluded by means of the enrolment criteria.

In selecting the outcome measures we have attempted to satisfy the most recent guidelines for the evaluation of anti-dementia drugs^{17,18} and to choose measures with documented evidence for validity, reliability, and sensitivity to change. The NAA and SKT are adapted to the specific needs of elderly people. In this trial we used parallel versions. The analysis of repeated measurements revealed a rather instable pattern of change in both the Ginkgo and placebo subgroups, for each of the outcome measures (SKT, CGI, NAA). This may reflect a lack of reliability, but may also be considered an indication that the intervention effect apparently was too weak to overrule random measurement error.

4.2 COMPARABILITY

Selection bias at baseline was largely prevented through random allocation. Additional measures to enhance comparability included restriction of the study population by applying rigorous eligibility criteria, and prestratification for diagnostic subgroup (dementia vs AAMI) and change in cognitive status during the 6-month pretreatment period (improvement or no change vs deterioration). Several other variables were taken into account in our stratified and multivariate regression analyses. A comparison of various baseline characteristics showed only minor, chance-related differences between the intervention groups. The small numbers of dropouts during the entire 24-week intervention period were equally distributed over the initial Ginkgo and placebo groups.

To enhance the comparability of interventions we pursued standardisation of treatment conditions. Subjects with a recent history of potentially influential types of medical and social care were excluded beforehand. Any planned or spontaneously occurring co-interventions were registered and taken into consideration as part of the data analysis. 28 Study subjects violated the predefined criteria for compliance with the study medication, 14 during the first part and another 14 during the second part of the trial. The non-compliance rate was 14% (23/166) for the initial Ginkgo group (doses combined) and 10% (5/48) for the initial placebo group. During the first 12 weeks relevant changes in either co-medication or care setting were notified for 19% (31/166) of the initial Ginkgo users (doses combined) and 17% (8/48) of the initial placebo users. These proportions amounted to 42% (33/79) and 32% (14/44) over the entire 24-week intervention period for the subgroups of Ginkgo and placebo users who qualified for the confirmatory analysis. A per-protocol analysis, excluding all study participants with registered protocol

deviations of treatment or measurement, yielded the same results as the intention-to-treat analysis.

We are convinced that the outcomes were assessed under blind conditions. Checks after 4 and 18 weeks of intervention did not reveal an association between the actual and the perceived type of treatment for any of the outcome reviewers involved (patients, nursing staff, interviewers). The layout of the trial forced us to employ many different raters and interviewers, who conducted the screening, baseline, and follow-up measurements. By and large, they were well prepared and dedicated to their job. The study protocol prescribed that, in general, repeated measurements for each study subject had to be conducted by the same rater or interviewer. We noted only a limited number of deviations from the pre-scheduled measurement time table; these were mainly caused by practical circumstances (holidays, etc.).

Inevitably, the analysis of the study results was hampered by the occurrence of missing values. For instance, after 24 weeks 27 out of the 214 randomized subjects were left without one or more SKT subtest scores, 18 without CGI assessment, and 34 without one or more NAA item scores. Based on the 'last observation carried forward' principle combined with the 'mean value of valid items/subtests' principle, these missing values could be substituted for 24, 14, and 25 cases, respectively. The application of several other accepted methods of missing value replacement yielded comparable results. Moreover, the similarity between the results of the intention-to-treat and per-protocol analyses was striking, suggesting that the occurrence and the 'repair' of missing values had not seriously affected the Ginkgo effect estimation. In summary, we do not believe that the trial results are biased by differential misclassification. We realize, however, that non-differential misclassifica-

tion may have caused some degree of extinction of a real existing effect.

4.3 OTHER CONSIDERATIONS

Its multipurpose character made our study design less efficient for answering the main research question, dealing with the efficacy of Ginkgo. The confirmatory analysis dealt with only 123 out of the 214 randomized subjects, with quite an imbalance between the contrasted treatments. An excess of study participants was allocated to the initial Ginkgo groups to enable answering the question regarding the persistence of any effect after stopping the Ginkgo treatment. Nevertheless, the trial had enough statistical power to be able to detect a substantial treatment effect.

Still, the results of our trial are dissenting from the results of the vast majority of previously reported Ginkgo trials in the field of cerebral insufficiency and dementia. After the systematic review work of Kleijnen et al^{1,22} new evidence for a beneficial effect of Ginkgo was provided in several trials, most of them with a higher methodological quality and more focussed on dementia than the earlier trials²³⁻²⁹. Up to now, we identified 9 Ginkgo trials (probably) focussing on patients with clear-cut dementia; 8 on patients with an AAMI-like diagnosis of cognitive impairment; 11 on patients with non-cognitive symptoms of cerebral insufficiency (tinnitus, vertigo, headache, etc.); 5 on patients with cerebral insufficiency diagnosed by means of a more or less standardized checklist of cognitive and non-cognitive symptoms; 3 on patients with affective disorders, especially depressive mood, as the leading symptom of cerebral insufficiency; and 18 on patients with cerebral insufficiency, not clearly specified³⁰. A beneficial Ginkgo effect was reported from the overwhelming majority of these trials. In the review of Kleijnen et al, only 8 of the 40 studies were assessed to be

of 'acceptable' methodological quality. We have noted that most later Ginkgo trials fulfilled considerably higher methodological requirements, both regarding internal validity and study size. Nevertheless, recently Letzel et al concluded that only a few of the trials included in their review²¹, met the current standards for the diagnosis of dementia (mild to moderate cases, accepted diagnostic criteria) and for the demonstration of a therapeutic effect (three levels of efficacy and clinical global assessment)^{24,29}. In addition to an improved methodological quality, we also noted a tendency over the years for smaller effects to be reported.

We believe that the Maastricht Ginkgo Trial is one of the most rigorous studies ever conducted in this field, especially with regard to random allocation, placebo control, and blinded outcome measurement. Our trial results suggest that treatment with Ginkgo is not efficacious, irrespective of dose, in elderly patients with mild to moderate dementia or age-associated memory impairment. Following a Bayesian approach we tend to conclude that the Maastricht Ginkgo Trial has scaled down our belief in a beneficial effect of Ginkgo for dementia and memory impairment. We realize, however, that the negative results of our trial cannot fully neutralize the positive results of previous studies. We have not been able to refute the positive results of some recent trials in particular. These trials appeared to meet quite high methodological standards and had several design aspects in common with our own study. As yet, we cannot fully rule out the possibility that our trial has been an 'outlier by chance'.

Contributors

M van Dongen and E van Rossum were involved in the study design, data collection and analysis. H Sielhorst contributed to the data collection and the organisation of the field work. A Kessels supervised the random allocation procedure and was involved in the

data analysis. P Knipschild supervised the Maastricht Ginkgo Trial and was especially involved in designing the trial and interpreting the results.

Acknowledgments

We are indebted to A Versantvoort, E Douven, and M Moll for assistance during the data collection phase (screening interviews); to J Kleijnen for his role in initiating the Maastricht Ginkgo Trial; to S van de Crommert for assistance in data processing and analysis; to the patients and their relatives, as well as to the management and the nursing staff teams of the participating homes for the elderly, for their cooperation and involvement in data collection; to the interviewers appointed to the Interview Service Department of Maastricht University for their contributions to the data collection; and to M van de Westelaken of the Department of Clinical Pharmacology of the Maastricht Academic Hospital for advice on pharmacological matters. The study was sponsored by Dr. Willmar Schwabe Arzneimittel GmbH, Karlsruhe, Germany. Dr. R Hörr, Dr. M Kieser, and Dr. G Meng of the Schwabe company provided us with useful advices during all phases of the trial.

REFERENCES

1. Kleijnen J, Knipschild P. *Ginkgo biloba*. *Lancet* 1992; 340: 1136-9.
2. Cott J. NCDEU update. Natural product formulations available in Europe for psychotropic indications. *Psychopharmacol Bull* 1995; 31: 745-51.
3. Grünwald J. The European phytomedicines: market figures. *Herbalgram* 1995; 34: 60-5.
4. DeFeudis FV. *Ginkgo biloba extract (EGb 761). From chemistry to the clinic*. Wiesbaden: Ullstein Medical, 1998.
5. Van Dongen M, Van Rossum E, Kessels A, Sielhorst H, Knipschild P. *The Maastricht Ginkgo Trial. Final report*. Results of a randomized clinical trial of the efficacy of a standardized treatment with *Ginkgo biloba* extract (EGb 761) in elderly people with dementia (mild to moderate degree) or memory impairment. Maastricht: Maastricht University, Department of Epidemiology, 1998.
6. APA. *DSM-III-R. Diagnostic and Statistical Manual of Mental Disorders*. 3th ed. Washington, D.C.: American Psychiatric Association, 1987.
7. WHO. *ICD-10: Tenth revision of the International Classification of Diseases*. Chapter V (F): Mental and behavioural disorders (including disorders of psychological development). Clinical descriptions and diagnostic guidelines. Geneva: World Health Organization, 1991.
8. Crook TH, Larrabee GJ, Youngjohn JR. Diagnosis and assessment of age-associated memory impairment. *Clin Neuropharmacol* 1990; 13: S81-91.
9. Zaudig M. A new systematic method of measurement and diagnosis of "mild cognitive impairment" and dementia according to ICD-10 and DSM-III-R criteria. *Int Psychogeriatr* 1992; 4: 203-19.
10. Folstein MF, Folstein SE, McHugh PR. Mini mental state: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12: 189-98.
11. Hachinski VC, Illif LD, Zilkha E. Cerebral blood flow in dementia. *Arch Neurol* 1975; 32: 632-7.
12. Overall JE, Schaltenbrand R. The SKT neuropsychological test battery. *J Geriatr Psychiatry Neurol* 1992; 5: 220-7.
13. Crook TH, Feher EP, Larrabee GJ. Assessment of memory complaint in age-associated memory impairment: the MAC-Q. *Int Psychogeriatr* 1992; 4(2):1 65-76.
14. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS). Recent evidence and development of a shorter version. In: Brink TL, ed. *Clinical gerontology. A guide to assessment and intervention*. New York: Hawthorn Press, 1986: 165-73.
15. Burke WJ, Roccaforte WH, Wengel SP. The short form of the Geriatric Depression Scale: a comparison with the 30-item form. *J Geriatr Psychiatry Neurol* 1991; 4: 173-8.

16. Amaducci L, Angst J, Bech P, Benkert O, Bruinvels J, Engel RR. Consensus conference on the methodology of clinical trials of 'nootropics', Munich, June 1989. Report of the Consensus Committee. *Pharmacopsychiatry* 1990; 23: 171-5.
17. CPMP. Note for guidance: antidementia medicinal products. CPMP Working Party on Efficacy of Medicinal Products. III/3705-91-EN, Draft 5. Brussels: Commission of the European Communities, 1992.
18. Menges K. Proof of efficacy of nootropics for the indication "Dementia" (Phase III) - Recommendations. *Pharmacopsychiat* 1992; 25: 126-35.
19. Oswald WD, Fleischmann UM, eds. *Nürnberger Alters-Inventar (NAI)*. NAI-Testmanual und -Text-band; 3th ed. Göttingen: Hogrefe, 1995.
20. Kieser M, Reitmeir P, Wassmer G. Test procedures for clinical trials with multiple endpoints. In: Vollmar J, ed. *Testing principles in clinical and preclinical trials. Biometrie in der chemisch pharmazeutischen Industrie* 6. Stuttgart: Gustav Fisher Verlag, 1995.
21. Letzel H, Haan J, Feil WB. Nootropics: efficacy and tolerability of products from three active substance classes. *J Drug Dev Clin Pract* 1996; 8:77-94.
22. Kleijnen J, Knipschild P. *Ginkgo biloba* for cerebral insufficiency. *Br J Clin Pharmacol* 1992; 34: 352-8.
23. Hofferberth B. The efficacy of EGb 761 in patients with senile dementia of the Alzheimer type, a double-blind, placebo-controlled study on different levels of investigation. *Human Psychopharmacology* 1994; 9: 215-22.
24. 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.
25. Vesper J, Hänsgen K-D. Efficacy of *Ginkgo biloba* in 90 outpatients with cerebral insufficiency caused by old age. Results of a placebo-controlled double-blind trial. *Phytomedicine* 1994; 1: 9-16.
26. Oswald WD, Hörr R, Oswald B, Steger W, Sappa J. Zur Verbesserung fluiden, kognitiver Leistungen mit Ginkgo-biloba-Spezial-extrakt EGb 761® bei Patienten mit leichten bis mittelschweren Hirnleistungsstörungen im Alter. *Zeitschrift für Gerontopsychologie und -psychiatrie* 1997; 10: 133-46.
27. Le Bars PL, Katz MM, Berman N, et al. A placebo-controlled, double-blind, randomized trial of an extract of *Ginkgo biloba* for dementia. *JAMA* 1997; 278: 1327-32.
28. Maurer K, Ihl R, Dierks T, Frölich L. Clinical efficacy of *Ginkgo biloba* special extract EGb 761 in dementia of the Alzheimer type. *J Psychiatr Res* 1997; 31: 645-55.
29. Haase J, Halama P, Hörr R. Wirksamkeit kurzdauernder Infusionsbehandlungen mit Ginkgo-biloba-Spezialextrakt EGb 761 bei Demenz vom vaskulären und Alzheimer-Typ. *Z Gerontol Geriatr* 1996; 29: 302-9.
30. Van Dongen MCJM, Van Rossum E, Knipschild PG. Efficacy of Ginkgo biloba extracts: evidence from randomized clinical trials. In: Van Beek TA, ed. *Ginkgo biloba*. Amsterdam: Harwood Academic Publishers, 1999 (*In Press*).

5

The efficacy of Ginkgo for elderly people with dementia and age-associated memory impairment: new results of a clinical experiment

Martien CJM van Dongen, MSc¹

Erik van Rossum, PhD¹

Alphons GH Kessels, MD¹

Hilde JG Sielhorst, RN^{1, 2}

Paul G Knipschild, MD, PhD²

¹ Department of Epidemiology, Maastricht University

² Department of General Practice, Maastricht University

ABSTRACT

Objective

To evaluate the efficacy, the dose-dependence, and the durability of the effect of the *Ginkgo biloba* special extract EGb 761 ('Ginkgo') in elderly people with dementia and age-associated memory impairment.

Design

A 24-week randomized, placebo-controlled trial. Patients with either mild to moderate dementia or age-associated memory impairment, were recruited in 39 homes for the elderly in the Netherlands. After giving informed consent, 214 eligible patients were randomized to three treatment groups: 240 mg Ginkgo ('high dose'), 160 mg Ginkgo ('usual dose'), or 0 mg Ginkgo (placebo) per day. By the end of 12 weeks of therapy the initial Ginkgo users were randomized once again, to either continued Ginkgo treatment or placebo treatment; initial placebo use was always prolonged for another 12 weeks. Outcomes were assessed after 12 and 24 weeks of intervention. These included neuropsychological testing (trail making speed (NAI-ZVT-G), digit memory span (NAI-ZN-G), and verbal learning (NAI-WL)), clinical assessment (presence and severity of geriatric symptoms (SCAG), depressive mood (GDS), self-perceived health and memory status (report marks)), and behavioural assessment (self-reported level of daily life activities (NAI-NAA)).

Results

An intention-to-treat analysis showed no effect on each of the outcome measures for participants who were assigned to Ginkgo (n=79) compared to placebo (n=44) for 24 weeks. After 12 weeks of treatment the groups with 'high dose' or 'usual dose' Ginkgo (n=166) performed slightly better with regard to self-reported activities of daily life, but slightly worse with regard to self-perceived health status, compared to placebo (n=48). No beneficial effects of a higher dose or a prolonged duration of Ginkgo treatment were found. We could not detect any subgroup that benefitted from Ginkgo. Ginkgo use was also not associated with the occurrence of (serious) adverse events.

Conclusion

Ginkgo does not seem to be effective in the treatment of elderly people with mild to moderate dementia or age-associated memory impairment. Our results contrast strongly with those of earlier Ginkgo trials.

1 INTRODUCTION

Ginkgo biloba special extracts, such as EGb 761 and LI 1370, are obtained through a standardized process of enrichment and purification of the original, full extract from dried leaves of the *Ginkgo biloba* tree. They contain fixed concentrations of two active substance classes: Ginkgo flavonol glycosides (16-25%) and terpene trilactones (usually 6%). Preparations based on these special extracts have entered the European drug market since 1965. Nowadays the use of 'Ginkgo' is widespread, either as an officially licensed drug that requires prescription by a physician, or as an over-the-counter medicine. Brand names like Tebonin[®], Kaveri[®], rōkan[®] and Tanakan[®] have become commercial bestsellers in countries such as Germany and France. In recent years the use of Ginkgo mono- or multipreparations seems to have expanded beyond its traditional boundaries, and Ginkgo products seem to be on the verge of conquering new markets^{1,2}.

Ginkgo drugs are prescribed and consumed for various indications. A recent review of the available scientific information concerning the *Ginkgo biloba* special extract EGb 761, which covers both the (bio)chemical and the clinical aspects, mentions four major types of action³. Firstly, Ginkgo is believed to exert a vasoregulatory action, which protects blood vessels and various tissues and might explain its use in managing cerebral insufficiency states, neurosensory disturbances and peripheral occlusive arterial disease. Secondly, it is supposed to have a cognition-enhancing action, which relates to its use in treating Alzheimer's disease and various other dementias. Thirdly, a stress-alleviating action is attributed to the use of Ginkgo, which might explain its claimed anxiolytic and antidepressive effects. And finally, a gene-regulatory action is suggested, which might explain why most clinically

beneficial effects of the extract seem to require repeated administration. A wealth of data from fundamental research is documented, which seem to refer to plausible mechanisms that underlie the proposed clinical actions attributed to Ginkgo constituents³.

Recent reviews of the clinical literature have strongly suggested a positive effect of Ginkgo in patients with cerebral insufficiency, peripheral arterial occlusive disease, and several other disorders⁴⁻¹². Especially the claimed beneficial influence of Ginkgo on cerebral insufficiency - including more focussed conditions, such as Alzheimer's disease, vascular dementia, and cognitive impairment no dementia - seems to attract the attention of the medical world increasingly. This interest is partly fed by the general experience that endeavours to develop either pharmacological or non-pharmacological treatments for dementia, which combine efficacy and safety, have not yet been very successful so far.

In 1993 we initiated a trial to confirm the results of previous studies of *Ginkgo biloba* in patients with cognitive decline. Moreover, our trial - the Maastricht Ginkgo Trial - aimed at investigating two additional research questions, regarding the dose-dependence and the persistence of the Ginkgo effect. We designed the trial in agreement with the latest guidelines for the evaluation of anti-dementia drugs and cognition enhancers^{13,14}. For a detailed description of the design, the conduct, and the results of the Maastricht Ginkgo Trial we refer to the final trial report¹⁵.

We already prepared a separate publication to communicate the results for some outcome measures used in our trial, viz. SKT (Syndrom Kurz Test, a brief test battery of memory and attention), CGI-2 (Clinical Global Impression of change in cognitive functioning), and NAI-NAA (Nürnberger Alters Inventar Alltagsaktivitäten Skala, an ac-

tivities of daily life scale, rated by senior caregivers)¹⁶. We applied in our trial several additional outcome measures, for various reasons: to provide for more subjectively oriented evidence of Ginkgo efficacy (patient self-assessment); to broaden the perspective from the cognitive domain to the affective domain (in particular depression); to elaborate on the effects demonstrated by the outcome measures referred to above, and to identify clinical effects linked up with specific cognitive processes and information-processing pathways; and, finally, to contest the results of previous trials, by applying identical, frequently used outcome measures (e.g., SCAG). The current publication deals with the results concerning these additional outcome parameters.

2 METHODS

2.1 PATIENT SELECTION

The target population encompassed subjects from two diagnostic categories: patients with dementia according to DSM-III-R¹⁷ and ICD-10 criteria¹⁸, and not-demented patients with age-associated memory impairment (AAMI)¹⁹⁻²¹. Patients with Alzheimer's disease (AD), vascular dementia (VD), or a mixed type of dementia were included in the dementia group.

The sources of recruitment were 39 old people's homes in the southern part of the Netherlands. The patient selection was based on a multi-stage screening procedure. For the differential diagnosis of dementia we applied the SIDAM^{22,23}, a Germany-based semi-structured interview, which entails among others the Mini-Mental State Examination (MMSE)²⁴ and the Hachinski Ischemic Score (HIS)²⁵. The SIDAM keeps close track of the DSM-III-R criteria for dementia. We admitted only subjects with a mild or moderate stage of dementia, as as-

essed by the results of the SIDAM interview and by a score between 8 and 23 on the Syndrom Kurz Test (SKT; Syndrome Short Test; full range: 0-27). The SKT is a brief psychometric test battery for the assessment of memory and attention^{26,27}.

For the AAMI diagnosis subjects had to satisfy the following inclusion criteria: impaired cognitive functioning, indicated by a score between 8 and 23 on the SKT; impaired cognitive functioning shown by the presence of self-reported memory complaints on a modified version of the Memory Assessment Clinics Questionnaire (MAC-Q), a brief, 5-item memory questionnaire referring to memory problems in daily life²⁸ (total score ≥ 12 (full range: 5-25), assigned by either the study subject or the nursing staff; and absence of dementia (SIDAM).

In addition to these inclusion criteria subjects of both subgroups had to meet the following enrolment criteria: age ≥ 50 years; written informed consent by patient and/or legal representative; absence of severe depression (Geriatric Depression Scale (GDS), 15-item version, total score < 11 (full range: 0-15))^{29,30}; presence of an adequate level of pre-morbid intelligence (IQ > 80 , global assessment); sufficient compliance and absence of placebo-response during the 3-week run-in period; no expectation of premature withdrawal; absence of serious co-morbidity, in particular pathological conditions which may be considered either nontreatable underlying causes of dementia and cognitive disorders, or sources of interference with the trial conduct (e.g., various neurological disorders, brain traumata, tumours, severe infectious diseases, absorption disorders); and absence of impermissible co-interventions, in particular drugs with a debilitating influence on psychical or cognitive functioning and drugs with a claimed nootropic action (e.g., antipsychotic drugs, antiparkinson medication, neuroleptics, cholinergic therapy, antidepressants, vasoactive drugs).

2.2 INTERVENTION

The experimental intervention consisted of a daily intake of either 160 mg or 240 mg Ginkgo. Two filmcoated tablets were administered per day, each containing either 80 mg or 120 mg of Ginkgo, (EGb 761, standardized on 24% Ginkgo flavonol glycosides and 6% terpene trilactones (Ginkgolides, bilobalid)). The placebo intervention consisted also of two tablets per day, that contained no Ginkgo extract. Special efforts were made to improve the comparability of verum and placebo, with respect to appearance, colour, smell, taste, granularity and solubility. In order to mimic the typical taste of Ginkgo extract, 2 ml quinine hydrochloride was added to the placebo tablets. The study participants took their daily dose at fixed times (morning and evening, during meals). Most of them did this under direct supervision of a nursing staff member. The study medication was packaged and labelled in accordance with the requirements of the Dutch Act of Medicine Supply and the guidelines for Good Clinical Practice.

2.3 RANDOM ALLOCATION

The trial was designed as a randomized, placebo-controlled, double-blind trial. In order to be able to answer all three research questions - addressing the efficacy, the dose-effect relationship, and the persistence of any effect of *Ginkgo biloba* treatment, respectively - we used a two-stage randomization strategy (see figure 1: study design). After a 3-week run-in period on placebo the eligible subjects were randomly allocated to one of three regimens: Ginkgo 761 160 mg/day ('usual dose'), Ginkgo 240 mg/day ('high dose'), or placebo. This tripartition aimed at the elucidation of the dose-dependence of any treatment effect. After having completed the first 12-week treatment period, subjects exposed to Ginkgo ini-

tially, were randomized once again to enter a second 12-week treatment period, either to continued Ginkgo or to placebo therapy; placebo use during the first 12 weeks was sustained for another 12 weeks. The purpose of the second randomization procedure was to enable an analysis of the durability of any intervention effect, after discontinuation of initial Ginkgo therapy.

A computer generated random sampling set was used for concealed random allocation of the study subjects. Random number assignment and labelling and distribution of the study medication were performed independently. Before randomization we prestratified the eligible study subjects according to their diagnostic subcategory (dementia vs AAMI) and perceived change in memory complaints during the 6-month pre-treatment phase (improvement or no change vs impairment). The first randomization step (week 0) was conducted in blocks of five patients and according to a 2 : 2 : 1 allocation ratio (Ginkgo 240 mg : Ginkgo 160 mg : placebo). To enhance the power of the study the original ratio was changed to 1 : 1 : 1 later on. The second randomization step (week 12) involved blocks of two patients and a 1 : 1 (Ginkgo : placebo) allocation ratio.

Eventually, this two-phasic random allocation strategy yielded an exposure contrast between five treatment groups of approximately equal size: 1. Ginkgo 240 mg (24 weeks); 2. Ginkgo 160 mg (24 weeks); 3. placebo (24 weeks); 4. Ginkgo 240 mg (12 weeks) + placebo (12 weeks); 5. Ginkgo 160 mg (12 weeks) + placebo (12 weeks).

2.4 OUTCOME MEASURES

The current guidelines for the evaluation of the efficacy of anti-dementia and cognition-enhancing drugs require that outcomes are measured at three different function levels: psychometric, psychopathological, and be-

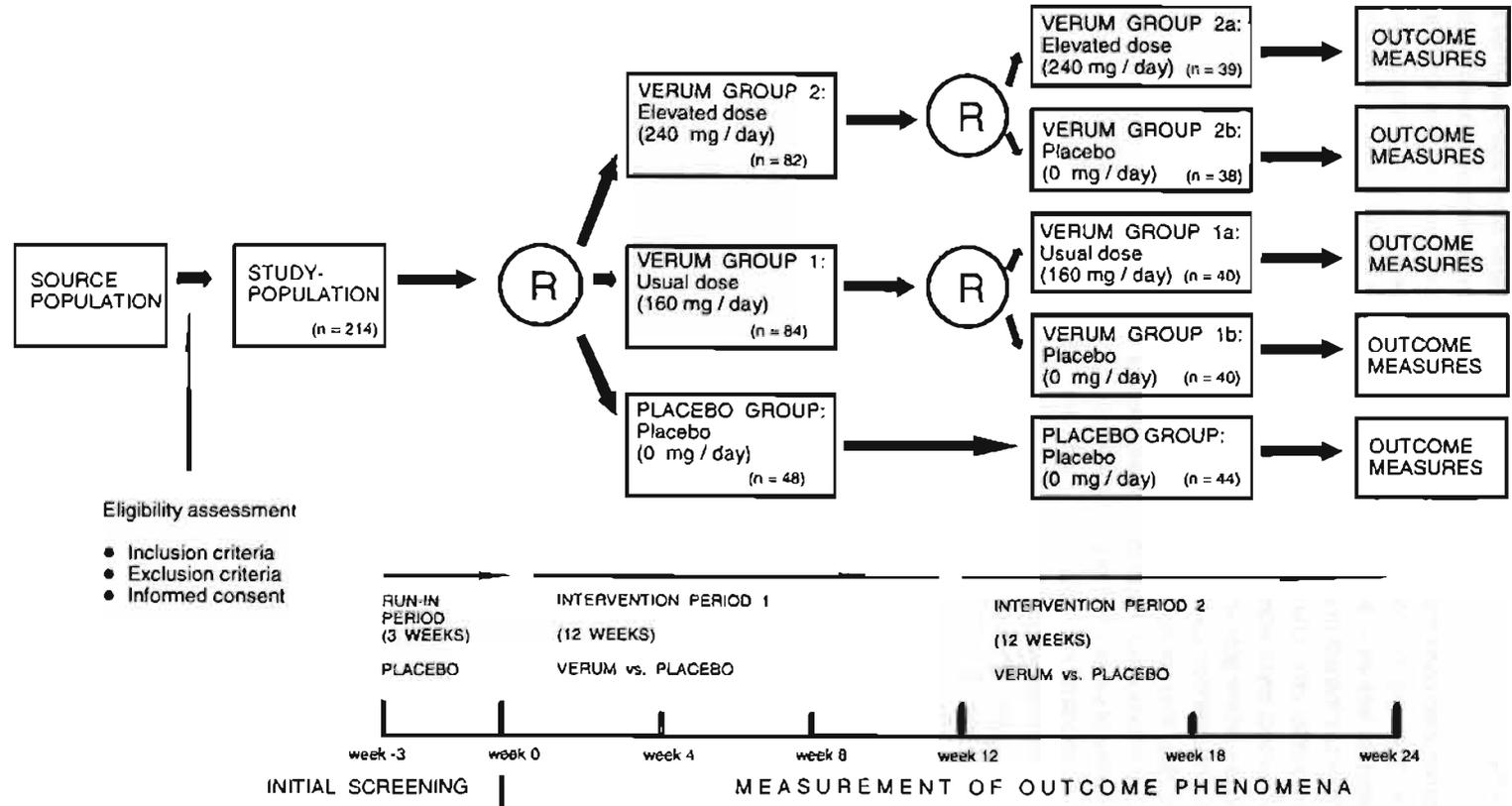


Figure 1. Design of the Maastricht Ginkgo Trial.

havioural functioning^{13,31}. As stated in the introduction, the confirmatory analysis of our trial results was based on three primary outcome measures that are not reported here. This article describes the results for several additional outcome parameters. These include 3 other psychometric tests, 4 measures of clinical assessment, and 1 measure related to behaviour:

- The *Zahlen-Nachsprechen-Test G* (ZN-G; Digit Memory Span G). This test, an adapted version of the classic digit span test, is a short term memory test aimed at the immediate reproduction of digit combinations of increasing length. The test entails two complementary formats: forward and backward reproduction. The total score (range: 0-17) is obtained by summation of the scaled scores for both formats.
- The *Wortliste* (WL; Word List). This is a verbal learning type of test, in which 8 carefully selected words have to be recalled and recognized. The scores obtained for the free, immediate recall procedure and the delayed recognition procedure (after about 20 minutes, mixed up with 8 distracter items) are combined to give a total score ranging from 0 to 16.
- The *Zahlen-Verbindungs-Test G* (ZVT-G). This is an adapted version of the classic trail making test, in which 30 numbers have to be connected in the right sequence and as quickly as possible. The test is supposed to measure planning, organization, and flexibility.
- The *Sandoz Clinical Assessment-Geriatric Scale* (SCAG). The SCAG is an 18-item geriatric symptom rating scale³²⁻³⁵. The target symptoms are supposed to represent several domains of impaired functioning: cognitive dysfunction, disturbed interpersonal relationship, affective disorders, inability to cope with the ordinary activities of daily life, and somatic dysfunction (e.g., fatigue, dizziness, poor

appetite). The total score ranges from 18 to 126, but is sometimes expressed as the mean item score.

- *Depressive mood*, expressed by the 15-item version of the Geriatric Depression Scale (GDS)^{29,30,36,37}. The GDS total score ranges from 0 to 15³⁸⁻⁴⁰.
- *Self-perceived health status*. This was measured on a 0-10 scale ('report mark').
- *Self-perceived memory status*. This was measured on a 0-10 scale as well ('report mark').
- The *Nürnberger Alters Alltagsaktivitäten Skala* (NAA; Nuremberg Gerontopsychological Rating Scale for Activities of Daily Living). The NAI-NAA has been designed to measure cognition-related behaviour. It consists of 20 items, covering instrumental daily life activities (e.g., getting dressed, performing light household tasks), social activities (e.g., visiting other people), and general cognitive performance (e.g., writing postcards, financial management). Summation of the individual item scores yields a total score from 20 ('very good') to 60 ('very poor'). Here we discuss the results of the self-assessment version of the NAA.

The ZVT-G, ZN-G, WL and NAA are all components of the *Nürnberger Alters Inventar* (NAI; Nuremberg Gerontopsychological Inventory), a German battery of psychometric tests, questionnaires and observational rating scales, that aims to measure cognitive functioning, performance of daily life activities, wellbeing and care needs. Although classic psychometric concepts underlie many of these tests and rating scales, most of them have been adapted to the specific needs of the elderly population⁴¹.

Two sources of information were used to collect data regarding these outcome measures: observation-based assessments by the nursing staff in the homes for the elderly (SCAG); and structured interviews (includ-

Table 1. Timing and scale ranges of outcome measures.

Measure	Time schedule (weeks)							Range of scores	
	-3	0	4	8	12	18	24	Status score	Change score
<i>Neuropsychological assessment</i>									
• NAI-ZVT (Zahlen-Verbindungs-Test: trail making test; seconds)		X			X		X	0 to <u>300</u>	-300 to <u>+300</u>
• NAI-ZN (Zahlen-Nachsprechen-Test: digit span test)		X			X		X	0 to <u>17</u>	-17 to <u>+17</u>
• NAI-WL (Wordliste: verbal learning test):									
- Forward		X			X		X	0 to <u>8</u>	-8 to <u>+8</u>
- Backward		X			X		X	0 to <u>8</u>	-8 to <u>+8</u>
- Total		X			X		X	0 to <u>16</u>	-16 to <u>+16</u>
<i>Clinical assessment</i>									
• Sandoz Clinical Assessment-Geriatric Scale (SCAG) (18 items, mean item score; hetero-assessment)		X	X	X	X	X	X	<u>1.0</u> to 7.0	-6.0 to <u>+6.0</u>
• SCAG-19 (overall impression; hetero-assessment)		X	X	X	X	X	X	<u>1</u> to 7	-6 to <u>+6</u>
• Geriatric Depression Scale (GDS-15 items; self-assessment)	X				X		X	<u>0</u> to 15	-15 to <u>+15</u>
• Health status (report mark; self-assessment)		X	X	X	X	X	X	0 to <u>10</u>	-10 to <u>+10</u>
• Memory status (report mark; self-assessment)		X	X	X	X	X	X	0 to <u>10</u>	-10 to <u>+10</u>
<i>Behavioural assessment</i>									
• NAI-NAA (Nürberger Alters Inventar - instrumental activities of daily life of elderly people; self-assessment)		X			X		X	<u>20</u> to 60	-40 to <u>+40</u>

Underlined printing is used to represent the most favourable extreme status score or change score

ing psychometric testing) by trained interviewers (the other outcome measures). Outcomes were measured at baseline and after 4, 8, 12, 18, and 24 weeks of treatment. Table 1 contains an overview of the measurement time schedule and the score ranges for each of the outcome measures presented here.

For continuous monitoring of the occurrence of (serious) adverse events we used a checklist of health complaints. This list was completed by the nursing staff at each measuring time point. Use of co-interventions and level of compliance with the study medication (pill counting) were registered as well.

2.5 DATA MANAGEMENT AND STATISTICAL ANALYSIS

Firstly, we performed an intention-to-treat analysis by comparing the 0-24 week changes in scores between the subgroups that received Ginkgo (160 mg and 240 mg

combined) and placebo during the entire 24-week treatment period. All study subjects for whom both baseline and 24-week follow-up scores could be made available, qualified for this analysis. Missing values were substituted according to a predefined replacement algorithm, with the last-observation-carried-forward principle as the main feature.

We also analysed the effects of 12 weeks of Ginkgo treatment, by comparing the 0-12 week changes in scores between the combined Ginkgo groups and the placebo group. To identify a possible dose-effect relationship we contrasted the changes in scores for each of the dosage categories (Ginkgo 240 mg/day, Ginkgo 160 mg/day, placebo), both after 24 weeks and 12 weeks of intervention. To assess the persistence of any intervention effect we compared the changes in scores after 24 weeks between the study subjects who were assigned to continue on Ginkgo, either 240 or 160 mg/day, and those

who were assigned to stop with Ginkgo after 12 weeks of treatment.

The statistical procedures included the calculation of the means and standard deviations of the changes in score on the outcome parameters for each of the comparison groups, as well as the calculation of the point estimates and the 90% confidence intervals of the mean change differences between the comparison groups. We performed multiple linear regression analyses to assess the adjusted effects for each of the outcome measures. Adjustment was made for baseline differences with regard to several potential confounders of the Ginkgo effect: age, gender, diagnostic subcategory (AAMI vs dementia), pretreatment level of change in memory complaints, baseline value of the outcome measure concerned, assessed quality of the old people's home, level of co-interventions during follow-up.

In addition to the intention-to-treat analysis we performed a per-protocol analysis. To qualify for the per-protocol analysis study subjects had to fulfil minimum requirements regarding predefined levels of compliance with the study medication ($\geq 75\%$ intake), adherence to the outcome measurement time table, and no exposure to unpermitted co-interventions.

3 RESULTS

The study participants were recruited from a total number of 4,459 residents who were present in the 39 homes for the elderly that cooperated in our trial. Of this population 781 were pointed out as potential study candidates by the senior caregivers, on suspicion of serious cognitive decline. From January 1994 to June 1996, 513 of them were actually screened to verify the presence of a state of mild or moderate dementia or cognitive impairment. Based on the screening results, a check of relevant medical and pharmaco-

therapeutical data, and their tentative consent, 256 subjects were admitted to a 3-week placebo run-in period preceding the random allocation and actual start of the intervention. By the end of this period, 214 subjects could be randomized: 63 demented and 151 non-demented patients with substantial cognitive decline. Over the next 24 weeks of follow-up 18 subjects withdrew prematurely, 13 during the first, and another 5 during the second 12-week intervention period. The dropouts were distributed quite homogeneously over the three initial treatment groups: 4/48 (8%), 5/84 (6%), and 9/82 (11%) for placebo, Ginkgo 160 mg, Ginkgo 240 mg, respectively.

Table 2 demonstrates the absence of marked differences in baseline characteristics between the three initially formed treatment groups. Probably the most striking difference was found for the trail making test (ZVT-G). It should be noted, however, that this measure was characterized by some outliers and a rather skewed distribution of individual scores, making the mean and standard deviation less suitable as indicators for central tendency and spread.

Tables 3 and 4 show the changes in the outcome measures for placebo and Ginkgo ('high' dose and 'usual' dose combined), after 24 weeks (table 3) and 12 weeks (table 4) of uninterrupted treatment. The tables include effect estimates both before and after adjustment for confounding. After 24 weeks of treatment statistically significant effects in favour of Ginkgo were found only for the self-perceived level of activities of daily life (NAA) and for trail making speed (ZVT-G); these effects were already present after 12 weeks of treatment. However, both effects were far from impressive and disappeared after adjustment for confounding, except for the 12 week difference in NAA-score. Furthermore, we found a statistically significant effect - in favour of placebo - for self-

Table 2. Distribution of baseline characteristics, by intervention subgroup at the start of the trial. Absolute numbers (percentages) for categorical variables, and mean values (standard deviations) for (semi-)continuous variables.

		placebo (n=48)	Ginkgo 160 mg (n=84)	Ginkgo 240 mg (n=82)
Gender	male	9 (19%)	13 (16%)	12 (15%)
	female	39 (81%)	71 (85%)	70 (85%)
Living situation	alone	43 (90%)	73 (87%)	75 (92%)
	together	5 (10%)	11 (13%)	7 (9%)
Cognitive impairment (assessed by staff)	mild	22 (46%)	37 (44%)	35 (43%)
	moderate	26 (54%)	47 (56%)	47 (57%)
Perceived change in memory complaints over last 6 months	increased much	6 (13%)	7 (9%)	9 (11%)
	increased	12 (25%)	25 (31%)	24 (29%)
	the same	27 (56%)	48 (59%)	46 (56%)
	decreased	3 (6%)	1 (1%)	3 (4%)
	decreased much	0 (0%)	0 (0%)	0 (0%)
State of dementia AAMI	dementia	14 (29%)	25 (30%)	24 (29%)
		34 (71%)	59 (70%)	58 (71%)
Age		82.6 (5.7)	82.8 (4.9)	83.3 (5.3)
Quefelet Index		25.0 (3.7)	24.2 (3.8)	24.9 (3.4)
MMSE (Mini-Mental State)		18.5 (4.5)	18.3 (4.8)	18.4 (4.2)
MAC-Q, modified, assessed by patient (subjective memory complaints)		12.6 (2.7)	12.7 (3.3)	12.0 (2.9)
MAC-Q, modified, assessed by nursing staff (subjective memory complaints)		15.2 (4.4)	16.0 (4.8)	16.1 (3.5)
NAI-ZVT-G (trail making test)		106.8 (57.6)	139.8 (77.4)	122.8 (76.1)
NAI-ZN-G (digit span)		8.1 (1.9)	8.0 (1.5)	8.0 (1.5)
NAI-WL-tot (word list; total score)		5.2 (3.1)	4.8 (3.4)	4.8 (3.0)
SCAG (geriatric symptoms; mean score)		2.5 (0.9)	2.6 (1.0)	2.7 (1.0)
GDS (depression)		3.0 (2.0)	2.5 (2.2)	2.5 (2.4)
Health status (report mark)		7.2 (1.5)	7.4 (1.2)	7.3 (1.4)
Memory status (report mark)		6.7 (1.3)	7.1 (1.1)	6.9 (1.2)
NAI-NAA (activities of daily life; self-assessment)		37.0 (6.4)	37.8 (6.1)	37.1 (5.3)

Scale ranges (underlined printing indicates the most favourable extreme):

MMSE: 0-30; MAC-Q: 5-25; NAI-ZVT-G: 0-300; NAI-ZN-G: 0-17; NAI-WL-tot: 0-16; SCAG: 1.0-7.0; GDS: 0-15;

health status: 0-10; memory status: 0-10; NAI-NAA: 20-60

perceived memory status after 12 weeks. Correction for confounding suppressed this effect, whereas it produced a new statistically significant effect for self-perceived health status. In general, we did not find any remarkable shift in score on most of the outcome parameters during the course of the intervention, neither in the Ginkgo group nor in the placebo group.

Given the absence of a positive effect for both Ginkgo dosage groups combined, a dose-related gradient seemed to be very unlikely in advance. Only at two occasions 'high dose' Ginkgo produced a stronger effect than 'usual dose' Ginkgo, once in a positive (self-perceived health status), and once in a negative sense (WL). After adjustment for confounding both effects were not statisti-

Table 3. Comparison of the mean baseline scores, mean 24-week follow-up scores, unadjusted mean 0-24 w. changes in score, and adjusted mean 0-24 w. changes in score (regression coefficients) on the outcome variables, for the placebo group (n=44) and the combined Ginkgo groups (n=79).

Variable	Intervention subgroup	Mean 0 w. baseline score	Mean 24 w. follow-up score	Change in score over 24 weeks mean (SD)	Unadjusted / Adjusted * 0-24 week change difference: mean (90% CI)
trail making speed (NAI-ZVT-G)	placebo	103.1	115.5	-12.5 (56.2)	+23.5 (+4.7; +42.2)
	Ginkgo	128.3	117.3	11.0 (60.4)	+18.4 (-0.7; +37.5)
digit span (NAI-ZN-G)	placebo	8.09	7.89	-0.20 (1.68)	-0.08 (-0.60; +0.44)
	Ginkgo	8.06	7.79	-0.28 (1.64)	-0.01 (-0.54; +0.52)
word list, total score (NAI-WL-tot)	placebo	5.19	4.19	-1.00 (2.86)	+0.40 (-0.52; +1.31)
	Ginkgo	4.24	3.64	-0.60 (2.88)	+0.12 (-0.69; +0.93)
geriatric symptoms, mean (SCAG)	placebo	2.48	2.45	0.03 (0.74)	+0.04 (-0.20; +0.29)
	Ginkgo	2.75	2.68	0.07 (0.81)	+0.01 (-0.22; +0.24)
depressive mood (GDS)	placebo	3.00	4.14	-1.14 (3.14)	+0.55 (-0.37; +1.47)
	Ginkgo	2.60	3.18	-0.58 (2.82)	+0.54 (-0.46; +1.54)
health status (report mark)	placebo	7.14	7.30	0.16 (1.29)	-0.32 (-0.77; +0.12)
	Ginkgo	7.43	7.27	-0.16 (1.49)	-0.15 (-0.56; +0.26)
memory status (report mark)	placebo	6.66	6.91	0.25 (1.14)	-0.28 (-0.64; +0.09)
	Ginkgo	6.99	6.96	-0.03 (1.19)	-0.15 (-0.51; +0.21)
activities of daily life (NAI-NAA)	placebo	37.00	39.57	-2.57 (4.61)	+1.65 (+0.20; +3.09)
	Ginkgo	37.82	38.74	-0.92 (4.63)	+1.32 (-0.13; +2.77)

* Adjusted change difference printed in italics. The number of subjects involved in the regression analysis varied from 118 (NAI-ZVT-G) to 122 (SCAG, report mark health status).

Scale ranges (underlined printing indicates the most favourable extreme):

status scores: NAI-ZVT-G: 0-300; NAI-ZN-G: 0-17; NAI-WL-tot: 0-16; SCAG: 1-0-7.0; GDS: 0-15; health status: 0-10; memory status: 0-10; NAI-NAA: 20-60

change scores: NAI-ZVT-G: -300 - +300; NAI-ZN-G: -17 - +17; NAI-WL-tot: -16 - +16; SCAG: -6.0 - +6.0; GDS: -15 - +15; health status: -10 - +10; memory status: -10 - +10; NAI-NAA: -40 - +40

cally significant anymore. This is demonstrated by the data in table 5.

Likewise, no significant beneficial effects were found for continued in comparison with discontinued use of *Ginkgo biloba* (see table 6).

To evaluate the possibility that a beneficial effect of Ginkgo might be the privilege of one or more subgroups of the study population, we performed a series of stratified analyses. These highlight the treatment effects for vari-

ous subcategories within the study population, based on the following baseline characteristics: age, diagnostic category (dementia vs AAMI), MMSE-score, SKT baseline score, BLE-score (number of experienced biological life events with a possible influence on cognitive functioning). All these baseline characteristics have been dichotomized. A qualitative summary of the results of these analyses is presented in table 7. In this table for each subgroup and each out-

Table 4. Comparison of the mean baseline scores, mean 12-week follow-up scores, unadjusted mean 0-12 w. changes in score, and adjusted mean 0-12 w. changes in score (regression coefficients) on the outcome variables, for the placebo group (n=48) and the combined Ginkgo groups (n=166).

Variable	Intervention subgroup	Mean 0 w. baseline score	Mean 12 w. follow-up score	Change in score over 12 weeks: mean (SD)	Unadjusted / Adjusted * 0-12 week change difference: mean (90% CI)
trail making speed (NAI-ZVT-G)	placebo	102.4	117.0	-14.6 (59.1)	+20.6 (+3.8; +37.3)
	Ginkgo	130.1	124.1	6.0 (58.3)	+14.5 (-2.1; +31.1)
digit span (NAI-ZN-G)	placebo	8.11	8.16	0.05 (1.18)	-0.05 (-0.47; +0.37)
	Ginkgo	8.06	8.05	-0.01 (1.57)	-0.03 (-0.43; +0.37)
word list, total score (NAI-WL-tot)	placebo	5.36	5.21	-0.14 (2.86)	-0.15 (-1.00; +0.70)
	Ginkgo	4.86	4.57	-0.29 (2.97)	-0.02 (-0.88; +0.84)
geriatric symptoms, mean (SCAG)	placebo	2.47	2.44	0.04 (0.61)	+0.15 (-0.06; +0.36)
	Ginkgo	2.68	2.50	0.18 (0.80)	+0.12 (-0.08; +0.32)
depressive mood (GDS)	placebo	3.05	4.25	-1.20 (3.63)	+0.59 (-0.27; +1.44)
	Ginkgo	2.52	3.14	-0.62 (2.84)	+0.60 (-0.31; +1.51)
health status (report mark)	placebo	7.21	7.26	0.04 (1.63)	-0.28 (-0.70; +0.13)
	Ginkgo	7.36	7.12	-0.24 (1.47)	-0.38 (-0.73; -0.04)
memory status (report mark)	placebo	6.72	7.00	0.28 (1.00)	-0.38 (-0.71; -0.04)
	Ginkgo	7.01	6.92	-0.09 (1.27)	-0.25 (-0.55; +0.05)
activities of daily life (NAI-NAA)	placebo	36.64	38.41	-1.77 (5.02)	+1.35 (+0.19; +2.51)
	Ginkgo	37.48	37.90	-0.42 (3.80)	+1.35 (+0.22; +2.49)

* Adjusted change difference printed in italics. The number of subjects involved in the regression analysis varied from 189 (NAI-WL-tot) to 206 (SCAG)

Scale ranges (underlined printing indicates the most favourable extreme):

status scores: NAI-ZVT-G: 0-300; NAI-ZN-G: 0-17; NAI-WL-tot: 0-16; SCAG: 1-0-7.0; GDS: 0-15; health status: 0-10; memory status: 0-10; NAI-NAA: 20-60

change scores: NAI-ZVT-G: -300 - +300; NAI-ZN-G: -17 - +17; NAI-WL-tot: -16 - +16; SCAG: -6.0 - +6.0; GDS: -15 - +15; health status: -10 - +10; memory status: -10 - +10; NAI-NAA: -40 - +40

come parameter a positive 24-week treatment effect in favour of Ginkgo is indicated by a '+' sign, whereas a positive effect in favour of placebo is indicated by a '-' sign; each effect that attained statistical significance ($\alpha = 0,10$) is marked by 's'. Bold printing of the corresponding '+' or '-' mark indi-

cates the subgroup with the most beneficial result.

Table 7 suggests that actually none of the evaluated subgroups consistently benefitted from Ginkgo (compared with placebo).

Table 5. Comparison of unadjusted and adjusted mean changes in score over 24 weeks on outcome measures, Ginkgo high dosage (240 mg, 0-24 weeks; n=40) vs Ginkgo usual dosage (160 mg, 0-24 weeks; n=39) vs placebo (0 mg, 0-24 weeks; n=44).

Variable	Intervention subgroup	Number	0-24 w. change in score:		Unadjusted / Adjusted * 24 week change difference, Ginkgo 240 vs 160: mean (90% CI)		
			mean	(SD)			
Trail making speed (NAI-ZVT-G)	placebo	42	-12.48	(56.24)			
	Ginkgo 160	40	14.24	(64.29)			
	Ginkgo 240	37	7.45	(56.63)	-6.8	(-29.8 ; +16.2)	
						-2.2	(-27.0 ; +22.6)
Digit span (NAI-ZN-G)	placebo	44	-0.20	(1.68)			
	Ginkgo 160	40	-0.18	(1.78)			
	Ginkgo 240	38	-0.39	(1.50)	-0.22	(-0.84 ; +0.40)	
						-0.01	(-0.70 ; +0.68)
Word list, total score (NAI-WL-tot)	placebo	42	-1.00	(2.86)			
	Ginkgo 160	40	0.03	(2.64)			
	Ginkgo 240	38	-1.26	(3.02)	-1.29	(-2.36 ; -0.22)	
						-1.37	(-2.78 ; +0.04)
Geriatric symptoms (SCAG; mean score)	placebo	44	0.03	(0.74)			
	Ginkgo 160	40	0.13	(0.91)			
	Ginkgo 240	39	0.02	(0.69)	-0.11	(-0.41 ; +0.19)	
						-0.23	(-0.54 ; +0.08)
Depressive mood (GDS)	placebo	44	-1.14	(3.14)			
	Ginkgo 160	40	-0.57	(2.93)			
	Ginkgo 240	37	-0.59	(2.74)	-0.02	(-1.10 ; +1.06)	
						-0.10	(-1.32 ; +1.12)
Health status (report mark)	placebo	44	+0.16	(1.29)			
	Ginkgo 160	40	-0.53	(1.24)			
	Ginkgo 240	39	+0.21	(1.64)	+0.73	(+0.19 ; +1.27)	
						+0.41	(-0.15 ; +0.97)
Memory status (report mark)	placebo	44	0.25	(1.14)			
	Ginkgo 160	39	-0.18	(1.02)			
	Ginkgo 240	39	0.13	(1.34)	+0.31	(-0.14 ; +0.76)	
						+0.23	(-0.26 ; +0.72)
Activities of daily life (NAI-NAA)	placebo	44	-2.57	(4.61)			
	Ginkgo 160	40	-0.43	(4.41)			
	Ginkgo 240	38	-1.45	(4.85)	-1.02	(-2.77 ; +0.73)	
						-1.43	(-3.21 ; +0.35)

* Adjusted change difference printed in italics.

Scale ranges (underlined printing indicates the most favourable extreme):

status scores: NAI-ZVT-G: 0-300; NAI-ZN-G: 0-17; NAI-WL-tot: 0-16; SCAG: 1.0-7.0; GDS: 0-15; health status: 0-10; memory status: 0-10; NAI-NAA: 20-60change scores: NAI-ZVT-G: -300 - +300; NAI-ZN-G: -17 - +17; NAI-WL-tot: -16 - +16; SCAG: - 6.0 - +6.0; GDS: -15 - +15; health status: -10 - +10; memory status: -10 - +10; NAI-NAA: -40 - +40

Table 6. Comparison of unadjusted and adjusted mean changes in score over 24 weeks on the outcome measures, Ginkgo 160/240 (24 weeks) (n=79) vs Ginkgo 160/240 (12 weeks) + placebo (12 weeks) (n=78).

Variable	Intervention subgroup	Number	0-24 w. change in score: mean (SD)		Unadjusted / Adjusted* 24 week change difference: mean (90% CI)	
Trail making speed (NAI-ZVT-G)	Ginkgo + placebo	74	2.6	(54.3)	<i>+8.3</i>	<i>(-7.0 ; +23.8)</i>
	Ginkgo + Ginkgo	77	11.0	(60.4)		
Digit span (NAI-ZN-G)	Ginkgo + placebo	76	-0.07	(1.63)	<i>+10.7</i>	<i>(-4.3 ; +25.7)</i>
	Ginkgo + Ginkgo	78	-0.28	(1.64)		
Word list, total score (NAI-WL-tot)	Ginkgo + placebo	73	-1.25	(2.62)	<i>-0.22</i>	<i>(-0.65 ; +0.22)</i>
	Ginkgo + Ginkgo	78	-0.60	(2.88)		
Geriatric symptoms (SCAG, mean score)	Ginkgo + placebo	78	0.12	(0.77)	<i>-0.06</i>	<i>(-0.48 ; +0.36)</i>
	Ginkgo + Ginkgo	79	0.07	(0.81)		
Depressive mood (GDS)	Ginkgo + placebo	77	-0.69	(2.97)	<i>+0.64</i>	<i>(-0.10 ; +1.39)</i>
	Ginkgo + Ginkgo	77	-0.58	(2.82)		
Health status (report mark)	Ginkgo + placebo	76	-0.17	(1.43)	<i>+0.70</i>	<i>(-0.21 ; +1.61)</i>
	Ginkgo + Ginkgo	79	-0.16	(1.49)		
Memory status (report mark)	Ginkgo + placebo	75	-0.12	(1.40)	<i>-0.05</i>	<i>(-0.25 ; +0.16)</i>
	Ginkgo + Ginkgo	78	-0.03	(1.19)		
Activities of daily life (NAI-NAA)	Ginkgo + placebo	77	-1.18	(4.35)	<i>-0.02</i>	<i>(-0.22 ; +0.40)</i>
	Ginkgo + Ginkgo	78	-0.92	(4.63)		

* Adjusted change difference printed in italics.

Scale ranges (underlined printing indicates the most favourable extreme):

status scores: NAI-ZVT-G: 0-300; NAI-ZN-G: 0-17; NAI-WL-tot: 0-16; SCAG: 10-70; GDS: 0-15; health status: 0-10; memory status: 0-10; NAI-NAA: 20-60

change scores: NAI-ZVT-G: -300 - +300; NAI-ZN-G: -17 - +17; NAI-WL-tot: -16 - +16; SCAG: -6.0 - +6.0; GDS: -15 - +15; health status: -10 - +10; memory status: -10 - +10; NAI-NAA: -40 - +40

4 DISCUSSION

Our trial did not reveal any systematic and clinically meaningful effect of Ginkgo on any of the outcome measures. Statistically significant effects regarding the time needed to complete the trail making test (NAI-ZVT-G) and the self-assessed level of activities of daily life (NAI-NAA) were detected initially, but these effects were partly extinguished after the influence of baseline differences was

taken into account through multiple linear regression analysis. Therefore, the current results seem to confirm the results reported previously for the SKT (Syndrom Kurz Test), the CGI-2 (Clinical Global Impression), and the NAI-NAA (Nürnbenger Alters Inventar Alltagsaktivitäten Skala) applied by the nursing staff, that also failed to demonstrate any beneficial effect of Ginkgo¹⁶.

Table 7. Evaluation of the efficacy of Ginkgo treatment (240 and 160 mg/day combined) in comparison with placebo treatment, after 24 weeks of intervention, for various subgroups (qualitative presentation of adjusted effects).

Outcome after 24 weeks	Baseline characteristic									
	Age (years)		Diagnosis		MMSE-score		SKT (baseline)		BLE-score	
	>83 (n=53)	≤83 (n=70)	Dem. (n=36)	AAMI (n=87)	<19 (n=66)	≥19 (n=57)	≥16 (n=57)	<16 (n=66)	≥2 (n=54)	<2 (n=69)
NAI-ZVT-G	+ s	+	+	+	+	+ s	+	+	-	+
NAI-ZN-G	-	+	+	+	+	-	+	-	+	-
NAI-WL-ir	+	-	+	+	+	+	+ s	-	+	+ s
NAI-WL-dr	+	-	+	-	+	-	+	- s	-	+
SCAG	+	-	-	+	-	+	-	+	-	-
GDS	-	+	+	+	+	+	-	+	+	+
Health status	-	-	-	+	+	-	+	-	+	-
Memory status	-	-	-	-	+	-	-	+	+	-
NAI-NAA	+	+	+	+	+	+	+	+	+ s	+

+ = Ginkgo better than placebo; - = placebo better than Ginkgo; 0 = no difference between Ginkgo and placebo; s = statistically significant effect, in favour of either Ginkgo or placebo; bold + enlarged sign (+, -, 0) indicates subgroup with the best effect of Ginkgo treatment; MMSE = Mini-Mental State Examination (0 - 30), measured during screening; SKT = Syndrom Kurz Test (memory and attention; Q - 27); BLE-score = Number of biological life events (Q - 5), measured as part of baseline interview)

Our trial is out of tune with the vast majority of placebo-controlled randomized trials conducted to assess the clinical efficacy of Ginkgo in patients with various types of cerebral insufficiency. A recent review of Ginkgo and cerebral insufficiency included 55 trials, 43 of which addressed cognition-related phenomena⁴². Nine trials, mostly of quite recent date, focussed on patients with dementia, either Alzheimer's disease or vascular dementia. They all showed positive, often statistically significant effects of Ginkgo compared with placebo. Eight trials dealt with patients suffering from cognitive impairment, probably not being dementia. For these trials positive effects of Ginkgo were reported as well, although not unequivocally and only for some of the outcome measures. Five trials included patients diagnosed by means of a symptom checklist that listed a more or less standardized combination of

cognitive and non-cognitive signs of cerebral insufficiency. Patients treated with Ginkgo were found to be significantly better off in each of these trials. Three trials had their primary focus on affective sequelae of cerebral function impairment, in particular depression. Ginkgo treatment caused alleviation of the state of depressive mood in each trial. Finally, 18 placebo-controlled trials dealt with patients diagnosed with a rather diffuse, unspecified form of cerebral insufficiency. Again, most of the authors reported a beneficial effect according to most of the endpoints measured⁴².

A trail making (number-connection) test in some form has been applied as one of the endpoints in seven previous placebo-controlled trials with Ginkgo. Statistically significant results were found in three of them, after 6-12 weeks of intervention⁴³⁻⁴⁵. A positive trend was found in four trials that made

use of the NAI-ZVT test⁴⁶⁻⁴⁹. The mean changes in ZVT-score (seconds) in our trial were +6 vs -15 after 12 weeks, and +11 vs -13 after 24 weeks (Ginkgo vs placebo). The change differences were not statistically significant anymore after adjustment for confounding.

Digit span testing, although not by means of the NAI-ZN subtest, was performed in five earlier trials⁵⁰⁻⁵⁴. The changes in our trial were small and statistically not significant: -0.0 vs +0.1 after 12 weeks, and -0.3 vs -0.2 after 24 weeks.

Five Ginkgo trials incorporated the SCAG as one of the outcome measures, either in its original format^{51,55-57}, or in a slightly modified version⁵⁸. In four trials, with daily dosages varying from 80 to 160 mg, Ginkgo turned out to be superior to placebo after 3 months^{51,55-57}, in the fifth trial after 12 months of intervention⁵⁸. The mean changes in SCAG total score in our trial amounted to +3.2 vs +0.7 after 12 weeks, and to +1.3 vs +0.5 after 24 weeks.

The GDS was used to evaluate the effect of Ginkgo on depressive mood in one earlier placebo-controlled trial. However, no results were reported from this trial⁵⁹. In other trials alternative depression scales were used, e.g., the Hamilton Depression Scale (range: 1-17) in two trials, with positive results^{38,39}; Zung's Self-Rating Depression Scale (SDS) also in two trials^{40,60}; and the Montgomery Asberg Depression Rating Scale⁶¹. In our trial we failed to identify any significant changes in depressive mood with the GDS: -0.6 vs -1.2 after 12 weeks, and -0.6 vs -1.1 after 24 weeks.

Two earlier trials measured the influence of Ginkgo on the level of daily activities by means of the NAI-NAA rating scale. Haase et al reported a positive effect of Ginkgo after 4 weeks of intravenous administration⁶⁰, whereas Oswald et al found no effect⁴¹. Our trial came up with the following changes: -0.4

vs -1.8 after 12 weeks, and -0.9 vs -2.6 after 24 weeks of therapy.

Outcome measures comparable with the word list test and the report marks for health and memory were not used in any of the previous Ginkgo trials.

In our trial we recruited residents of old people's homes, who suffered from either dementia or age-associated memory impairment, using a multiphasic screening procedure. This screening procedure entailed a combination of diagnostic tools (questionnaires, psychometric tests, rating scales) which have proven to be valid and reliable. We preferred to include residents of homes for the elderly above elderly people still living at home for several reasons: to create a suitable sampling frame, harbouring many eligible study candidates; to be able to promote and supervise the level of compliance with the study medication; to facilitate the conduct of baseline and follow-up measurements as well as other logistic procedures; to gain easy access to additional (medical) information concerning the study subjects; and to avoid an overrepresentation of 'worriers', who tend to exhibit an unwarranted amount of fear regarding cognitive deterioration. We realize that, as a consequence of this choice, we have attracted a study population with a typical composition as to age, co-morbidity, et cetera. The living environment and care provided to these people may protect them to some extent from the mental challenges of daily life in the open society. This circumstance may have affected the correct classification of some of the study participants as demented or just memory impaired, as well some of the outcome measurements, especially those relying on self-assessment by the participants.

As far as the study treatment is concerned, we applied a high-quality *Ginkgo biloba* special extract (EGb 761), with a significant ex-

posure contrast between verum (either 160 or 240 mg of active substance per day) and placebo. We spent much effort to develop a valid placebo. A strict distribution schedule, with intake supervision by the nursing staff, and periodical pill counts, was used to enhance the compliance with the intervention regimen. Some potential non-compliers were identified and excluded already during the run-in phase, before randomization. Unwanted interference with co-interventions was precluded by means of the enrolment criteria.

In selecting the outcome measures we have attempted to satisfy the prevailing guidelines for the evaluation of anti-dementia drugs^{13,14}. Especially the NAI tests and rating scales, and the GDS were adapted to the specific need of elderly people. At consecutive occasions we applied parallel versions of NAI neuropsychological tests, to avoid repeated testing artefacts.

We took various measures to improve the comparability of the trial groups, the interventions to be evaluated, and the outcome assessments. To preclude any disturbing influences of prognostic subgroup differences, we randomly allocated the participants and also relied on restriction, prestratification, stratified analysis and multiple regression analysis. Only minor differences in baseline characteristics between the intervention groups were found. The number of dropouts was limited and quite homogeneously distributed among the intervention subgroups. In order to release the active Ginkgo compounds as the only source of intervention contrast, people exposed to a broad range of pharmacological and nonpharmacological cognition-affecting principles were filtered out and excluded before the start of the trial. We monitored relevant changes in drug use and medical care, as well as the occurrence of important psychological life events, continuously during the trial. Insufficient compli-

ance with the study medication, often in combination with premature withdrawal, was registered for 10%, 11%, and 17% of the initial placebo, Ginkgo 160 mg, and Ginkgo 240 mg groups, respectively. A per-protocol analysis, restricted to the 'valid cases' with regard to compliance, exposure to co-interventions, and timing of the outcome measures, did not change the conclusions of the intention-to-treat analysis.

The study outcomes were assessed under blind conditions. To verify blindness the patients themselves, the nursing staff and the interviewers involved in outcome measurements were all invited to report their study medication (Ginkgo 240 mg, Ginkgo 160 mg, placebo). These checks did not disclose any association between actual and perceived type of treatment, neither at 4 nor at 18 weeks after the start of the intervention. Given these observations we can hardly imagine that the trial results have seriously been distorted due to differential misclassification. However, we cannot totally rule out the possibility of some random outcome measurement error, that may have given rise to suppression of actual Ginkgo effects.

We conclude that our trial has failed to reproduce the beneficial effects of Ginkgo in elderly patients with dementia and age-associated memory impairment, which were demonstrated by many previous trials. The findings regarding the outcome parameters reported in this article, support the earlier conclusions of our trial report. In our opinion the absence of a consistent treatment effect cannot be attributed to flaws and fallacies embedded in the trial design, conduct and analysis. Following a Bayesian approach we feel that the results of the current trial have negatively affected the belief in a beneficial influence of Ginkgo on cognitive functioning.

Acknowledgements

We are indebted to A Versantvoort, E Douven, and M Moll for their assistance during the data collection phase (screening interviews); to J Kleijnen for his role in initiating the Maastricht Ginkgo Trial; to S van de Crommert for her assistance in the data processing and analysis; to the patients and their relatives, as well as to the management and the nursing staff teams of the participating homes for the elderly, for their cooperation and involvement in data collection; to the interviewers appointed to the Interview Service Department of Maastricht University for their contributions to the data collection; and to M van de Westelaken of the Department of Clinical Pharmacology of the Maastricht Academic Hospital for advice on pharmacological matters.

The study was sponsored by Dr. Willmar Schwabe Arzneimittel GmbH, Karlsruhe, Germany. Dr. R Hörr, Dr. M Kieser, and Dr. G Meng of Schwabe company at Karlsruhe provided us with useful advices during all phases of the trial.

REFERENCES

- Cott J. NCDEU update. Natural product formulations available in Europe for psychotropic indications. *Psychopharmacol Bull* 1995; 31: 745-51.
- Grünwald J. The European phytomedicines: market figures. *Herbalgram* 1995; 34: 60-5.
- DeFeudis FV. *Ginkgo biloba extract (EGb 761). From chemistry to the clinic*. Wiesbaden: Ullstein Medical, 1998.
- Weiß H, Kallischnigg G. Ginkgo-biloba-Extrakt (EGb 761). Meta-Analyse von Studien zum Nachweis der therapeutischen Wirksamkeit bei Hirnleistungsstörungen bzw. arterieller Verschlusskrankheit. *Munch Med Wschr* 1991; 133: 138-42.
- Kleijnen J, Knipschild P. Ginkgo biloba for cerebral insufficiency. *Br J Clin Pharmacol* 1992; 34: 352-8.
- Kleijnen J, Knipschild P. Ginkgo biloba. *Lancet* 1992; 340: 1136-9.
- Schneider B. Ginkgo-biloba-Extrakt bei peripheren arteriellen Verschlusskrankheiten. Meta-Analyse von kontrollierten klinischen Studien. *Arzneimittelforschung* 1992; 42: 428-36.
- Hopfenmüller W. Nachweis der therapeutischen Wirksamkeit eines Ginkgo biloba Spezialextraktes. Meta-Analyse von 11 klinischen Studien bei Patienten mit Hirnleistungsstörungen im Alter. *Arzneimittelforschung* 1994; 44: 1005-13.
- Letzel H, Haan J, Feil WB. Nootropics: efficacy and tolerability of products from three active substance classes. *J Drug Dev Clin Pract* 1996; 8: 77-94.
- Ernst E. Ginkgo biloba in der Behandlung der Claudicatio intermittens. Eine systematische Recherche anhand kontrollierter Studien in der Literatur. *Fortschr Med* 1996; 114: 85-7.
- Kanowski S. Ginkgo-biloba-Spezialextrakt. Nachgewiesene Wirksamkeit im Indikationsbereich Demenz. *Munch Med Wochenschr* 1997; 139: 39-42.
- Schulz V, Hubner WD, Ploch M. Clinical trials with phytopsychotherapeutic agents. *Phytomedicine* 1997; 4: 379-87.
- Menges K. Proof of efficacy of nootropics for the indication "Dementia" (Phase III) - Recommendations. *Pharmacopsychiatry* 1992; 25: 126-35.
- CPMP. *Note for guidance: antidementia medicinal products*. CPMP Working Party on Efficacy of Medicinal Products. III/3705-91-EN, Draft 5. Brussels: Commission of the European Communities, 1992.
- Van Dongen M, Van Rossum E, Kessels A, Sielhorst H, Knipschild P. *The Maastricht Ginkgo Trial*. Final report. Results of a randomized clinical trial of the efficacy of a standardized treatment with *Ginkgo biloba* extract (EGb 761) in elderly people with dementia (mild to moderate degree) or memory impairment. Maastricht: Maastricht University, Department of Epidemiology, 1998.
- Van Dongen MCJM, Van Rossum E, Kessels APH, Sielhorst HJG, Knipschild PG. Ginkgo for elderly people with dementia and age-associated

- memory impairment. A randomized clinical trial. Submitted for publication, 1999.
17. APA. *DSM-III-R. Diagnostic and Statistical Manual of Mental Disorders*. 3th ed. Washington, DC: American Psychiatric Association, 1987.
 18. WHO. *ICD-10: Tenth revision of the International Classification of Diseases*. Chapter V (F): Mental and behavioural disorders (including disorders of psychological development). Clinical descriptions and diagnostic guidelines. Geneva: World Health Organization, 1991.
 19. Crook TH, Bartus RT, Ferris SH, Whitehouse P, Cohen GD, Gershon S. Age-associated memory impairment: proposed diagnostic criteria and measure of clinical change - report of a National Institute of Mental Health working group. *Dev Neuropsych* 1986; 2: 261-76.
 20. Crook T, Bahar H, Sudilovsky A. Age-associated memory impairment: diagnostic criteria and treatment strategies. *Int J Neurol* 1987; 88: 21-22.
 21. Crook TH, Larrabee GJ, Youngjohn JR. Diagnosis and assessment of age-associated memory impairment. *Clin Neuropharmacol* 1990; 13: S81-91.
 22. Zaudig M, Mittelhammer J, Hiller W, et al. SIDAM - A structured interview for the diagnosis of dementia of the Alzheimer type, multi-infarct dementia and dementias of other aetiology according to ICD-10 and DSM-III-R. *Psychol Med* 1991; 21: 225-36.
 23. Zaudig M. A new systematic method of measurement and diagnosis of "mild cognitive impairment" and dementia according to ICD-10 and DSM-III-R criteria. *Int Psychogeriatr* 1992; 4: 203-19.
 24. Folstein MF, Folstein SE, McHugh PR. Mini mental state: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12: 189-98.
 25. Hachinski VC, Illif LD, Zilkha E. Cerebral blood flow in dementia. *Arch Neurol* 1975; 32: 632-7.
 26. Kim YS, Nibbelink DW, Overall JE. Factor structure and scoring of the SKT test battery. *J Clin Psychol* 1993; 49: 61-71.
 27. Overall JE, Schaltenbrand R. The SKT neuropsychological test battery. *J Geriatr Psychiatry Neurol* 1992; 5: 220-7.
 28. Crook TH, Feher EP, Larrabee GJ. Assessment of memory complaint in age-associated memory impairment: the MAC-Q. *Int Psychogeriatr* 1992; 4: 165-76.
 29. Sheik JI, Yesavage JA. Geriatric Depression Scale (GDS). Recent evidence and development of a shorter version. In: Brink TL, ed. *Clinical gerontology. A guide to assessment and intervention*. New York: Hawthorn Press, 1986: 165-73.
 30. Burke WJ, Roccaforte WH, Wengel SP. The short form of the Geriatric Depression Scale: a comparison with the 30-item form. *J Geriatr Psychiatry Neurol* 1991; 4: 173-8.
 31. CPMP. *Note for guidance on medicinal products in the treatment of Alzheimer's disease*. London: The European Agency for the Evaluation of Medicinal Products, Human Medicines Evaluation Unit, Committee for Proprietary Medicinal Products, 1997.
 32. Shader RI, Harmatz JS, Salzman C. A new scale for clinical assessment in geriatric populations: Sandoz Clinical Assessment-Geriatric (SCAG). *J Am Geriatr Soc* 1974; 22: 107-13.
 32. Shader RI, Harmatz JS, Salzman C. Sandoz Clinical Assessment-Geriatric (SCAG). *Psychopharmacol Bull* 1988; 24: 765-9.
 34. Venn RD. The Sandoz Clinical Assessment-Geriatric (SCAG) scale. A general-purpose psycho-geriatric rating scale. *Gerontology* 1983; 29: 185-98.
 35. Patin JR, Hamot HB, Singer JM. Replicated evidence on the construct validity of the SCAG (Sandoz Clinical Assessment-Geriatric) scale. *Prog Neuropsychopharmacol Biol Psychiatry* 1984; 8: 293-306.
 36. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1982; 17: 37-49.
 37. Sheikh JI, Yesavage JA, Brooks JO, et al. Proposed factor structure of the Geriatric Depression Scale. *Int Psychogeriatr* 1991; 3: 23-8.

38. Stocksmeier U, Eberlein M. Depressive Verstimmung bei Hirnleistungsstörungen. Wirkung eines Ginkgo-biloba-Extraktes in Doppelblind-Studie überprüft. *TW Neurologie Psychiatrie* 1992; 6: 74-76.
39. Schubert H, Halama P. Primärtherapieresistente depressive Verstimmung älterer Patienten mit Hirnleistungsstörungen: Wirksamkeit der Kombination von Ginkgo-biloba-Extrakt EGb 761 mit Antidepressiva. *Geriatric Forschung* 1993; 3: 45-53.
40. Halama P. Was leistet der Spezialextrakt EGb 761? Therapie mit *Ginkgo biloba* bei Patienten mit cerebrovaskulärer Insuffizienz und therapieresistenter depressiver Symptomatik. Ergebnisse einer placebokontrollierten, randomisierten doppelblinden Pilotstudie. *Therapiewoche* 1990; 40: 3760-5.
41. Oswald WD, Fleischmann UM, eds. *Nürnberger-Alters-Inventar (NAI)*. NAI-Testmanual und -Textband. 3th ed. Göttingen: Hogrefe, 1995.
42. Van Dongen M, Van Rossum E, Knipschild P. Efficacy of *Ginkgo biloba* extracts: evidence from randomized clinical trials. In: Van Beek TA, ed. *Ginkgo biloba*. Amsterdam: Harwood Academic Publishers, 1999 (*In press*).
43. Hofferberth B. Einfluß von *Ginkgo biloba* Extrakt auf neurophysiologische und psychometrische Meßergebnisse bei Patienten mit hirnganischem Psychosyndrom. Eine Doppelblindstudie gegen Plazebo. *Arzneimittelforschung* 1989; 39: 918-22.
44. Brüchert E, Heinrich SE, Ruf-Kohler P. Wirksamkeit von LI 1370 bei älteren Patienten mit Hirnleistungsschwäche. Multizentrische Doppelblindstudie des Fachverbandes Deutscher Allgemeinärzte. *Münch Med Wochenschr* 1991; 133 (Suppl. 30): S9 - S14.
45. Hofferberth B. Ginkgo-biloba-Spezialextrakt bei Patienten mit hirnganischem Psychosyndrom. Prüfung der Wirksamkeit mit neurophysiologischen und psychometrischen Methoden. *Münch Med Wochenschr* 1991; 133 (Suppl. 30): S30 - S33.
46. Hartmann A, Frick M. Wirkung eines Ginkgo-Spezial-Extraktes auf psychometrische Parameter bei Patienten mit vaskulär bedingter Demenz. *Münch Med Wochenschr* 1991; 133 (Suppl. 30): S23 - S25.
47. Halama P. *Ginkgo biloba*. Wirksamkeit eines Spezialextraktes bei Patienten mit zerebraler Insuffizienz. *Münch Med Wochenschr* 1991; 133: 190-4.
48. Maurer K, Ihl R, Dierks T, Frölich L. Clinical efficacy of *Ginkgo biloba* special extract EGb 761 in dementia of the Alzheimer type. *J Psychiatr Res* 1997; 31: 645-55.
49. Oswald WD, Hörr R, Oswald B, Steger W, Sappa J. Zur Verbesserung fluider, kognitiver Leistungen mit Ginkgo-biloba-Spezialextrakt EGb 761® bei Patienten mit leichten bis mittelschweren Hirnleistungsstörungen im Alter. *Zeitschrift für Gerontopsychologie und -psychiatrie* 1997; 10: 133-46.
50. Dieli G, La Mantia V, Saetta M, Costanzo E. Essai clinique à double insu du tanakan dans l'insuffisance cérébrale chronique. *Lavoro Neuropsychiatrico* 1981; 68: 1-7.
51. Weitbrecht WV, Jansen W. Doubleblind and comparative (*Ginkgo biloba* versus placebo) therapeutic study in geriatric patients with primary degenerative dementia - a preliminary evaluation. In: Agnoli J, Rapin R, Scapagnini V, Weitbrecht WV, eds. *Effects of Ginkgo biloba Extract on organic cerebral impairment*. Montrouge: John Libbey Eurotext Ltd., 1985: 91-99.
52. Wesnes K, Simmons D, Rook M, Simpson P. A double-blind placebo-controlled trial of Tanakan in the treatment of idiopathic cognitive impairment in the elderly. *Human Psychopharmacology* 1987; 2: 159-69.
53. Rai GS, Shovlin C, Wesnes KA. A double-blind, placebo controlled study of *Ginkgo biloba* extract ('tanakan') in elderly outpatients with mild to moderate memory impairment. *Curr Med Res Opin* 1991; 12: 350-5.
54. Maier-Hauff K. LI 1370 nach zerebraler Aneurysma-Operation. Wirksamkeit bei ambulanten Patienten mit Störungen der Hirnleistungsfähigkeit. *Münch Med Wochenschr* 1991; 133: S34 - S37.
55. Pidoux B, Bastien C, Niddam S. Clinical and quantitative EEG double-blind study of *Ginkgo bi-*

- loba* extract (GBE). *J Cereb Blood Flow Metab* 1983; 3: S556 - S557.
56. Halama P, Bartsch G, Meng G. Hirnleistungsstörungen vaskulärer Genese. Randomisierte Doppel-blindstudie zur Wirksamkeit von Ginkgo biloba Extrakt. *Fortschr Med* 1988; 106: 408-12.
57. Hofferberth B. The efficacy of EGb 761 in patients with senile dementia of the Alzheimer type; a double-blind, placebo-controlled study on different levels of investigation. *Human Psychopharmacology* 1994; 9: 215-22.
58. Taillandier J, Ammar A, Rabourdin JP, et al. Traitement des troubles du vieillissement cérébral par l'extrait de *Ginkgo biloba*. Étude longitudinale multicentrique à double insu face au placebo. *Presse Med* 1986; 15: 1583-7.
59. Israël L, Dell'Accio E, Martin G, Hugonot R. Extrait de *Ginkgo biloba* et exercices d'entraînement de la mémoire. Evaluation comparative chez des personnes âgées ambulatoires. *Psychologie Médicale* 1987; 19: 1431-9.
60. Haase J, Halama P, Hörr R. Wirksamkeit kurzdauernder Infusionsbehandlungen mit Ginkgo-biloba-Spezial extrakt EGb 761 bei Demenz vom vaskulären und Alzheimer-Typ. *Z Gerontol Geriatr* 1996; 29: 302-9.
61. 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.

6 GENERAL DISCUSSION

1 INTRODUCTION

This final chapter reflects on the findings regarding the efficacy of Ginkgo, which we presented in the previous chapters. Firstly, it addresses the results of our literature review of the controlled trials of Ginkgo in cerebral insufficiency and related disorders that have been published since the introduction of Ginkgo drugs in the mid-1960s. The rationale, the strategy and the results of this review were described in chapter 2. Secondly, it discusses the internal validity, the external validity, and the results of our own clinical trial - the Maastricht Ginkgo Trial. The design of this trial was described in chapter 3, the results in chapters 4 and 5. Finally, I will ascertain whether it is possible to reconcile the conflicting results of the Maastricht Ginkgo Trial and those of previous trials, and I will attempt to give my personal view on the efficacy of Ginkgo in patients with dementia, serious cognitive decline, and related disorders.

2 REVIEW OF PREVIOUS TRIALS

The Maastricht Ginkgo Trial was preceded by a considerable number of earlier trials, which all tried to demonstrate the efficacy of Ginkgo for cerebral insufficiency and related

disorders. In the recent past several attempts have been made to summarize the published evidence for the clinical efficacy and safety of Ginkgo. As we have explained in chapter 2, we tend to characterize most of them as 'narrative reviews', although some relied on a structured presentation and description of the main design and outcome characteristics of the trials included in the review¹⁻⁷. In addition to these reviews we came across two more systematically organised reviews, viz. a criteria-based systematic review, published by Kleijnen & Knipschild in 1992⁸ and a quantitative meta-analysis based on the statistical pooling of individual trial results, published by Hopfenmüller in 1994⁹.

The various reviewers unanimously arrived at the summary judgment that Ginkgo is efficacious in patients suffering from cerebral insufficiency. However, the authors of the criteria-based systematic review concluded that only 8 of the 40 trials showed a reasonable methodological quality, according to a predefined checklist of criteria. These trials were all published between 1986 and 1991, and all reported a positive effect in favour of Ginkgo⁸. To enable pooling Hopfenmüller restricted his meta-analysis to a subset of trials that all applied the *Ginkgo biloba* special extract LI 1370 (Kaveri[®] forte) as the active

intervention substance, and a standardized checklist of subjective, clinical symptoms of cerebral insufficiency as the main outcome measure⁹. Out of 11 trials which were considered eligible, only 8 qualified for the analyses based on the single symptoms of cerebral insufficiency and the clinical symptom total score; and only 6 trials could be used to perform a meta-analysis that highlighted the global assessment of the treatment effect, either by the physician or by the patient, as the outcome measure of interest.

In our review we have attempted to address all published Ginkgo trials that fulfilled minimum requirements with regard to methodology (randomization, comparison of Ginkgo with placebo or alternative treatment) and that focussed on any health problem covered by the nowadays somewhat obsolete concept of 'cerebral insufficiency'. For reasons stated in chapter 2 we have decided to abstain from both formal methodological quality assessment and statistical pooling of individual study or patient data. Any attempt to pool the relevant studies would be frustrated by the heterogeneity of study participants, interventions and outcomes. Probably the pooling operation performed by Hopfenmüller⁹ is all that can be achieved. It should be noted, however, that this meta-analysis dealt with the methodological quality of the studies only superficially, and that the Ginkgo trials published after 1994 - with a relatively high methodological quality - probably would not fit into the same pool. We have seriously considered an update or replay of the criteria-based systematic review conducted by Kleijnen and Knipschild⁸. But eventually we expected to encounter too many difficulties with the application of any of the existing criteria lists for assessing methodological quality, which have been designed according to the generally accepted principles of intervention research^{10,11} and which have been previously used in several formats and modifications in various do-

main of (para)medical intervention¹²⁻¹⁵. Especially the criteria referring to the composition of the study population and the adequacy of the outcome measures are difficult to apply for Ginkgo trials, as both the diagnosis of the assessment and the clinical course of dementia and related forms of cerebral insufficiency are surrounded by many uncertainties. The steadily evolving insights in this field are accompanied by periodical adaptations of the criteria and guidelines for diagnosis and outcome evaluation. Kleijnen & Knipschild have avoided this problem in their review by not specifying very strict and focussed quality criteria for study population and outcome measurement. Otherwise, many trials would probably have been excluded, or at least been assigned even lower quality scores on these aspects. This dilemma is also illustrated by the review published by Letzel et al⁵ in 1996, in which they compared the evidence for the efficacy and tolerability of Ginkgo and two other active substance classes. They concluded that only 2 of the 25 Ginkgo trials included in the review were in full agreement with the most up-to-date criteria for the evaluation of anti-dementia drugs, with regard to both patient definition (diagnosis of dementia) and outcome measurement. The application of other quality criteria appears to be more straightforward. Nevertheless, one should not trifle with them. For instance, while most Ginkgo trials are projected as placebo-controlled in the literature, the use of a placebo preparation may not automatically guarantee blinded outcome assessment. When preparing our trial we noticed that the development of a good placebo for Ginkgo is a difficult task, which apparently had not received too much attention¹⁶, even not in trials conducted in populations that are more familiar with Ginkgo. Poor reporting of a trial is another obstacle for sound methodological quality assessment.

The stronger the association between the methodological quality and the effect sizes of the studies, the more important it is to incorporate the variations in the quality into the estimation of the summary effect¹⁷. The fact that earlier reviews did not point to a strong association for Ginkgo trials made our decision to abandon formal quality assessment procedures easier.

Instead of statistical pooling or criteria-based reviewing, we have attempted to organize the relevant publications into several categories, according to their diagnostic orientation (dementia, age-associated memory impairment, clinical symptomatology, non-cognitive physical symptoms, non-cognitive affective symptoms, and non-specified cerebral insufficiency), followed by a detailed description of the design and results of the trials assigned to each category.

Being not too strict, we identified 9 trials that apparently focussed on demented patients. All but one showed positive effects of Ginkgo, for most of the applied outcome measures. However, especially for the most recent and most rigorously designed trials we found rather modest effect sizes. We detected 8 Ginkgo trials that included AAMI-like patients. Again, the results point at the superiority of Ginkgo. Most trial reports, however, showed positive results only for part of the outcome parameters evaluated. Most often the investigators had neglected the hierarchical ordering of these parameters. Eleven of 12 RCTs that dealt with non-cognitive complications of cerebral insufficiency, such as tinnitus, vertigo and hearing loss, reported positive, often statistically significant results. Very consistent results were also found for 5 trials that included patients with cognitive and non-cognitive subjective symptoms of cerebral insufficiency. In each trial a positive influence was found for almost all the evaluated symptoms. Furthermore, we identified 3 trials with patients suffering from depressive mood as the leading symp-

tom of cerebral insufficiency. All reported a positive effect of Ginkgo, which was statistically significant in two trials. Due to insufficient diagnostic information, we had to classify 18 trials into a residual category ('cerebral insufficiency unspecified'). Most of them dated back to the 1970s and 1980s. All reported a beneficial effect of Ginkgo for most of the outcome measures.

Our review seems to confirm the efficacy of Ginkgo for each of the clusters of trials which we have distinguished, with clustering being based on diagnostic subgroup. Nevertheless, we have identified also a few trials with negative results and quite a lot with inconclusive results for some of the outcome measures. For the trials that focussed on AAMI-like patients with cognitive impairment, the results were the least convincing. We could not attribute these inconsistencies to differences in research methods. Also, we have not received overt signals that publication bias might have flattered the picture of the efficacy of Ginkgo. The heterogeneity of the applied outcome measures prevented us from evaluating publication bias by plotting the effect sizes against the sample sizes of the individual trials ('funnel plot')¹⁸. We are aware that more subtle forms of publication bias, e.g., inconsistencies between protocolized and reported outcome measures, may also be responsible for a distorted image. Actually, this type of data manipulation can be unmasked and prevented only by setting up a register of relevant trials. Admission to the inception cohort of trials that make up the register, should be based on approval of the trial protocol, and the data presented in the final trial report should be checked with the original protocol. For the time being, however, we have no indication that discrepancies between protocol and published results have contaminated our review. We have refrained from applying formal methodological quality assessment procedures,

but we do believe that, on average, most recent trials harbour a relatively high amount of evidence, with respect to both methodological rigour and statistical power. We also have the impression that smaller effects are reported in the course of time, although time-related differences in the use of outcome measures preclude a direct comparison between trials that haven been conducted in different periods.

3 THE MAASTRICHT GINKGO TRIAL

Here we will first summarize the results of our trial and then we will discuss the probability that these results are spurious, because they have originated from incorrect design choices or dissimilarities in the trial. Finally, we will reflect briefly on the design and the results of the Maastricht Ginkgo Trial, and compare these with previous trials that addressed a similar research question.

3.1 SUMMARY OF THE RESULTS

The confirmatory analysis of our trial data was based on a statistical test, which integrated the changes in scores on the three primary outcome measures - a psychometric test of memory and attention (SKT), global impression of change in cognitive function (CGI-2), and instrumental activities of daily life, assessed by the nursing staff (NAA_{staff}) - over 24 weeks of treatment. This analysis did not show a statistically significant difference between the patients treated with Ginkgo and those treated with placebo during the entire intervention period.

For two of the three primary outcome measures (SKT, CGI-2) we found very small differences in change over 24 weeks, in favour of the Ginkgo group; these were neither statistically significant, nor clinically meaningful. For the third primary outcome measure

(NAA_{staff}) the difference was neglectable. Adjustment for possible confounding by other predictors of the relevant outcomes, did not yield substantial shifts of the estimated effects of Ginkgo treatment on each of the primary outcome measures. None of the differences approached statistical significance. The differences were also very small for most of the secondary outcome measures. However, for two of these measures - the level of instrumental activities of daily life, perceived by the participants themselves (NAA_{self}), and the trail making speed (ZVT) - the differences were statistically significant (< 0.05 , one-sided), on both occasions in favour of the Ginkgo group. Adjustment for confounding resulted in a reduction of the estimated effect sizes for most of the outcome parameters.

An additional responder analysis revealed that the numbers of participants who showed a clinically meaningful improvement for at least two of the three primary outcome measures, were almost equally distributed over the Ginkgo and the placebo subgroups, both after 24 weeks and after 12 weeks of intervention. This implies that, apart from the presence of any statistically significant effect, Ginkgo treatment did not yield any effects that we considered clinically relevant. The evaluation aimed at the elucidation of a possible dose-response relationship and the persistency of a possible intervention effect, did not show important differences, neither between the high-dose and low-dose, nor between the long-term and short-term Ginkgo users.

The results of the previous trials in this field justified a quite substantial belief in the efficacy of Ginkgo before the onset of the current trial. This belief motivated our decision to formulate separate study questions regarding a possible dose-response relation and the durability of any effect, in addition to the main research question, that referred to

the efficacy of Ginkgo as such. However, the current trial did not reveal any beneficial effect of Ginkgo. In the next section we will review possible explanations for these more or less unexpected, negative results.

3.2 REFLECTIONS ON THE DESIGN, CONDUCT AND RESULTS OF THE MAASTRICHT GINKGO TRIAL

First of all, we should consider the possibility that the validity of the trial has been affected by erroneous choices or flaws, which may have compromised the study results. Therefore, we will reflect briefly on the appropriateness of the Maastricht Ginkgo Trial to address the research question underlying the trial. This will be done by reviewing the problems that we have encountered regarding the following aspects: adequacy of study population, adequacy of intervention, adequacy of outcome measures, comparability of study participants (intervention subgroups), comparability of (co-)interventions, and comparability of outcome measurements. The choices with regard to the study population, the type of intervention, and the outcome measures relate mainly to the external validity, while systematic dissimilarities between treatment subgroups, interventions and outcome measurements refer to the internal validity.

3.2.1 Adequacy of study population

While selecting the study population, we came across several problems that may have influenced the study results. One of these concerned the limited amount of information available on the medical history of the candidate participants. Furthermore, it was not always easy to decide on the - conditional or unconditional - exclusion of candidate participants who violated the co-medication requirements. Sometimes such decisions tended to be somewhat arbitrary.

The decision to use old people's homes as the main source for recruitment of the study population had several far-reaching implications for the composition and characteristics of the study population, such as an overrepresentation of the very high age groups, a relatively short remaining life expectancy, and a rather high level of co-morbidity. The protective environment where these people live in, may make them less conscious and worried about their level of cognitive functioning. The care and shelter provided in the homes for the elderly apparently neutralize part of the mental challenges that are present in the 'open' community. The trial setting also complicated the correct classification of the study subjects as either 'demented' or just 'memory impaired'. The DSM-III-R criteria for dementia imply that the distinction between the two is based on the assessment of the influence of cognitive impairment at the level of social functioning. It appeared to be difficult to assess this influence under the prevailing sheltered conditions.

Although the delineation of the study population had several difficulties, we cannot imagine, however, that the absence of an intervention effect in our study can fully be attributed to the selection of our study population.

3.2.2 Adequacy of interventions

In the current trial we used tablets based on a high-quality standardized extract from dried leaves of the *Ginkgo biloba* tree (input : extract ratio = 50 : 1). The dosage schedule provided a daily exposure to relevant compounds that is considered to be crucial for proper action (80 or 120 mg of 'active' substance per tablet, containing 24% flavonol glycosides, 6% terpene trilactones; 2 tablets a day). A marked exposure contrast was generated between the verum groups - two modalities: a normal dose group, using two tablets containing 80 mg extract each day,

and a high dose group, using two tablets containing 120 mg extract each day - and the reference group (two tablets per day without any Ginkgo extract). Much effort was paid to the development of a valid placebo tablet.

We are confident that the Ginkgo and placebo tablets were manufactured properly. Packaging and labelling of the preparations was done in agreement with the Dutch Act of Medicine Supply and the Good Clinical Practice guidelines. During the trial the study medication was stored in accordance with the relevant prescriptions (e.g., environmental temperature), both at the study centre and in the participating old people's homes.

Each participant had to take two tablets a day (morning and evening, during the meal). This frequency of use should guarantee a continuous supply of the relevant body tissues with the active substances. The way of distribution of the study medication within each participating home for the elderly, and the supervision by the nursing staff can be seen as major safeguards for a sufficient level of compliance with the intervention regimen. In order to verify the level of compliance we have conducted pill counts at regular time intervals during the follow-up period. The run-in phase of the trial was used to identify potential non-compliers, who were excluded from the randomization procedure subsequently. Only a limited number of participants appeared to be self-supportive as to the use of medication. For some of them this evoked problems, as was illustrated by the pill count results. In one particular case no remaining tablets were returned after the first intervention period, while we counted a surplus of left-overs after the second intervention period. Apart from this incident we are not aware of major irregularities regarding the supply with study medication. We noticed only a limited amount of real non-compliance.

We consider it very unlikely that the absence of an intervention effect was caused by an inappropriate design and conduct of the study intervention.

3.2.3 Adequacy of outcome measurements

In designing the outcome measures for the Maastricht Ginkgo Trial we have tried to adhere to the most recent guidelines for the evaluation of anti-dementia drug trials. These guidelines prescribe that outcome measures must cover three domains of performance: cognitive functioning (assessment based on neuropsychological testing); psychopathology (clinical assessment); and behaviour (assessment based on behavioural observation or self-reporting). We have applied these guidelines to our situation by selecting a combination of tests, questionnaires and rating scales that we thought to be suitable for both their proven psychometric qualities (validity, precision, and sensitivity to change) and their level of adaptation to the requirements of the age group involved in the study. Nevertheless, it appeared not always easy to conduct the measurements properly, for varying reasons (e.g., exhaustion of the study subject; lack of motivation to expose oneself to the competitive atmosphere imposed by several of the performance tests). The noise introduced by this kind of measurement problems may have caused some decrease of the treatment effect. The scales and measurement schedules to be applied were pretested in the practical research setting, and, if necessary, modified on the basis of the experiences gained from these try-outs before the start of the trial.

The choice of the majority of the outcome measures was founded on documented evidence with regard to their clinimetric capacities. Unfortunately, the design and the planning of our study did not allow for extensive,

formal validation research during a pilot phase or as part of the actual intervention trial. We have not checked the inter-rater agreement of the various outcome measures, for instance. Therefore, we could not verify whether the documented properties of these psychometric tests and rating scales also held for our target population, with its specific composition (very old inhabitants of homes for the elderly), and for the measurement conditions that prevailed in our trial (measurement by a team of interviewers and by old people's home staff members). Face validity and content validity were taken into consideration during the selection of our measurement tools. Due to the lack of appropriate external criteria, the criterion validity could not be quantified.

An aspect of the reliability of the outcome measures that we were able to evaluate, was the stability of the measurement results and the effects that were registered over time. For this purpose we made an inventory of the number of subsequent 'one-way' changes in overall score for each of the three primary outcome measures (SKT, CGI-2 and NAA_{staff}). In doing this we considered every change in score between two consecutive measurement time points to be an effect, irrespective of its size, statistical significance and clinical meaningfulness. As such time-related scoring patterns are difficult to interpret for study subjects who shifted their medication in between, we will refer here only to the data for those subjects who underwent either Ginkgo or placebo treatment during the entire 24 week intervention period. The SKT-test has been issued at 6 different measurement time points, giving rise to 5 consecutive measurement intervals. Among the study population - 79 persons with continued Ginkgo treatment and 44 persons with continued placebo treatment - no cases with a consistent effect (either improvement, or no change, or impairment) at all 5 occasions could be identified. Two sub-

jects (1 Ginkgo 160 mg, 1 placebo) showed improvement, 1 subject (placebo) showed impairments and the status of 1 person (Ginkgo 240 mg) remained unchanged for 4 consecutive intervals. Six persons (2 Ginkgo 240 mg, 1 Ginkgo 160 mg, and 3 placebo) improved, and 1 person (Ginkgo 240 mg) deteriorated at 3 consecutive measurement points. All the other participants showed an even more fluctuating, in general non-monotonous response pattern. With regard to the CGI, which was scored at weeks 4, 8, 12, 18, and 24, we found a consistent positive change for 6 subjects (2, 4, and 0 for Ginkgo 240, Ginkgo 160, and placebo, respectively), a consistent negative change for another 6 subjects (2, 4, and 0), and a consistent 'no change' for 22 subjects (9, 8, and 5). And as far as the NAA_{staff} rating scale is concerned, 15 participants (5, 5, and 5) progressed, 20 (6, 7, and 7) regressed, and 3 (all Ginkgo 240 mg), remained unchanged both after 12 weeks and after 24 weeks of follow-up. These findings suggest that the stability of the treatment effect over time was rather low for the three primary outcome measures, and that episodes of deterioration or no change succeeded episodes of improvement. Obviously, the actual intervention effect was not impressive enough to overrule any possible random measurement errors.

The length of follow-up and the timing of the observations are other important aspects which determine the adequacy of outcome evaluation. We suppose that these aspects have been adequate in the current trial and have not precluded the detection of any relevant effect of the Ginkgo treatment.

We do not believe that the particular choice of the outcome measures in this trial can be held responsible for the absence of a treatment effect, although we cannot deny that the selection of measurement tools has been arbitrary to a certain extent, that we could

have used alternative options, and that feasibility considerations - e.g., the non-clinical setting of the trial - have influenced our decisions in this respect.

3.2.4 Comparability of intervention subgroups

Selection bias in the sense of a difference of prognosis between subjects allocated to the various treatment groups, seems very unlikely, given the nature of the study design (prospective cohort, random allocation to the intervention of interest). In an attempt to advance the between-group similarity we have applied several design measures. The restriction of the study population, guided by the set of exclusion criteria, can be seen as a first measure to guarantee the similarity of the intervention groups to be compared in this trial, in particular with regard to exposure to co-medication and co-morbidity before the start of the trial. Because we did not want to rely fully on the randomization procedure, we used prestratification before the random allocation procedure, for two important predictors of the disease prognosis: the severity of the disease state under study (dementia vs AAMI), and the estimated speed of deterioration of the cognitive function during the previous 6 months. In order to avoid a lack of balance over time, the eligible subjects were randomly allocated in blocks of five (first intervention period), respectively two (second intervention period).

We registered a limited number of drop-outs during the trial: 18 subjects over the entire 24-week period. These dropouts were distributed quite equally over the allocation subgroups.

We conclude that the comparability of the treatment groups in our trial has been satisfying. The existence of residual confounding cannot be ruled out completely, but it is very unlikely that any remaining confounding has caused spurious conclusions.

3.2.5 Comparability of (co-)interventions

Ideally, once prognostic similarity between the intervention groups has been established at baseline, all patients should be exposed to either the experimental or reference intervention under identical circumstances. Standardisation of the treatment conditions that prevail during the trial should be aimed at, in order to create comparable treatment conditions for all individual participants. Any remaining differences in co-intervention level should be controlled as part of the data analysis.

Subjects with cognition-related co-medication before the start of the trial were excluded from the randomization procedure. We also did not permit participants to engage in other activities aimed at cognitive enhancement, e.g., memory training courses. Changes in co-medication after the start of the trial were recorded. The same holds for other changes in care that might intervene with the effects of the study treatment or with a proper execution of the follow-up measurements (e.g., hospital or nursing home admission). We have attempted to adjust for these concurrent treatment influences by entering the level of co-interventions (present vs absent) as one of the independent variables in the multivariate regression models. Furthermore, we have registered the occurrence of important 'psychological life events' - life events that might affect the mental status, either positively or negatively. Since the general level of care and guidance provided in each home (attention paid to the residents, social activities organised, et cetera) may also influence the prognosis of cognitive function, we have paid special attention to this aspect: we classified the participating homes using a simple 'quality of care' index. This 'quick and dirty' judgement was added as a covariate in the regression models.

In addition to the identification of possible differences in the level of co-interventions, we evaluated also the degree of compliance with the intervention itself, i.e. the study medication. We identified 28 participants who did not meet the compliance criteria for the entire intervention period. 14 Of them violated these rules only during the second part of the trial. In most cases non-compliance with the prescribed study medication was induced by withdrawal from the trial. Five non-compliers were initially allocated to the placebo group (rate: 10.4%), and 23 to the combined verum groups (rate: 10.7%). We registered 39 persons with any form of co-intervention - either relevant change of co-medication, or any change in care - during the first intervention period (rates: 17.0% and 18.7% for the placebo group and the combined Ginkgo groups, respectively). Of the subjects who made up the placebo group and the combined Ginkgo group in the confirmatory analysis, 31.8% (14/44) and 41.8% (33/79) had been exposed to co-interventions at any time during the entire 24-week follow-up period, respectively. Nine participants, viz. 6.4% vs 3.6%, of the initial placebo and Ginkgo groups, experienced a change of co-medication at any time during the trial that may have influenced their cognitive state. Any change of care was recorded for 68 of the randomized patients: 12 (25.8%) in the placebo group and 56 in the combined Ginkgo groups (34.0%).

Although exposure to the experimental research conditions may have produced an upgrading of the general level of care and attendance within the participating homes, we assume that the placebo-induced blinding of all assessors has prevented the occurrence of any systematic differences between the intervention groups in this respect. A per-protocol ('valid cases') analysis, which excluded all participants with observed protocol deviations, either due to non-exposure to the study intervention or to not-allowed ex-

posure to co-interventions, yielded results that did not differ from the intention-to-treat analysis results.

We conclude that systematic differences in (co-)interventions between the study groups cannot be held responsible for the observed absence of a treatment effect.

3.2.6 Comparability of outcome measurements

In our trial the outcome measures were conducted under blind conditions. The results of a check on the thoroughness of the blinding procedure were satisfying. Therefore, we may assume that information bias caused by differential misclassification is not an issue of major concern. If any misclassification has occurred, this will have been nondifferential misclassification most probably, resulting in a partial extinction of the intervention effect. We involved many different interviewers and raters in the screening, baseline, and follow-up measurements. This variety of assessors may be regarded a weak point of the trial. In general we were confident, however, about their competences and capacities to perform their tasks properly. Moreover, they appeared to be very dedicated to their job. We did not have the opportunity to evaluate the inter- and intra-rater reliabilities of the measurement scales based on our own trial results. As a rule repeated measurements for each study subject were conducted by the same interviewer or nursing staff member. The baseline values of the SKT and GDS, however, were gathered by the members of the research team during the screening interviews, whereas the follow-up measurements were conducted by the team of interviewers.

We noted only a few deviations from the planned measurement time-table. Most often these were caused by practical circumstances (holidays, etc.).

The data collection procedures did not yield a complete data matrix. Whereas for the primary outcome measures the number of missing values was limited, we had to deal with more problems for some of the secondary outcome variables, e.g., depressive mood (GDS). Overseeing the available options, we gave priority to a rather simple and transparent replacement strategy to deal with missing values in the analysis. Substituted data were used to perform both the intention-to-treat analysis and the per-protocol analysis.

We consider it not very likely that incomplete measurement ('missing values') has influenced the study results to a considerable extent. The results of the per-protocol analysis almost equalled those of the intention-to-treat analysis. In addition to the preferred strategy for dealing with missing values, which rested on two principles - substitution by carrying forward the last valid observation, and substitution by the mean value of valid subtests - we examined the consequences of two other replacement strategies, but only for the primary outcome variables (SKT, CGI, NAA_{staff}). The first one implied that missing values at any measurement time point were estimated through single linear regression analysis, based on the valid scores for the same measure for the same individual. At least two valid repeated measurements had to be available. This procedure was conducted at the overall test level, not at the subtest or item level. The second additional strategy implied that missing values at any measurement time point were substituted by the worst measurement result ever for the same individual and the same measurement tool, over the entire 24-week intervention period. This substitution replaced the 'last observation carried forward' strategy and was combined with the mean value of valid subtests replacement strategy. It was done at the item or subtest

level. None of both additional replacement strategies urged us to revise our conclusions.

In summary, we consider it unlikely that incomparable outcome measurements have caused a substantial distortion of the real effect of Ginkgo treatment. However, some random measurement error may be responsible for an underestimation of an existing effect.

3.2.7 Other considerations

In addition to the main research question, that dealt with the efficacy of *Ginkgo biloba* treatment, we intended to answer two additional questions through the current trial, that addressed the persistence and the dose-relationship of any effect. Its 'multi-purpose' function made the design a little less efficient for answering the principal research question. Moreover, we did not manage to recruit the envisioned number of 300 subjects that we aimed at in advance. But even if we had accomplished this sample size, this would not have made the trial enough powerful at all to make the recorded effect sizes statistically significant. Several thousands of participants would have been needed to make the effect sizes found in our trial statistically significant.

3.3 ADVERSE EVENTS

In order to evaluate the tolerability of the study medication in our trial we registered the occurrence of ten common health complaints (predefined list), which we wanted to evaluate as potential side effects of the study medication, as well as the occurrence of serious adverse events (open reporting) and potential symptoms of addiction (predefined list of 7 symptoms).

Dizziness, nervousness and headache were relatively often reported as health com-

plaints during the trial. In almost all cases the nursing staff in the homes for the elderly tended not to attribute these complaints to the study medication. A statistical analysis did not reveal any association between type of study medication (Ginkgo vs placebo) and the occurrence of health complaints.

During the trial 35 serious adverse events were registered, which affected 33 participants (8 deaths, 25 hospital admissions, and 2 malignancies). For most of these events an association with the study medication was considered not probable or absent. A possible association was stated for only one event, a hospital admission after a stomach bleeding in a placebo user.

An inventory of symptoms indicative for a possible withdrawal effect was made after 18 and 24 weeks of intervention. Relatively high frequencies of three symptoms - lack of concentration, irritability, and fatigue - were reported for study subjects who had changed from Ginkgo to placebo after 12 weeks of intervention, compared with subjects who continued the use of Ginkgo tablets. The proportion of participants with at least one withdrawal symptom was about 2 times higher for the group that stopped compared to the group that continued with Ginkgo use, both after 18 and 24 weeks of follow-up. We found the same pattern for the three symptoms most closely relate to brain (dys)function (lack of concentration, irritability, fatigue). We could not find a reasonable explanation for these findings.

From these figures we conclude that the frequencies of non-specific health complaints and serious adverse events do not exceed the expected frequencies to a considerable extent, and that a clear relation with the type of study medication used cannot be established. It seems that the study medication appeared to be safe and highly tolerable.

4 COMPATIBILITY OF THE MAASTRICHT GINKGO TRIAL WITH PREVIOUS TRIALS

Most of the previous trials that evaluated the efficacy of Ginkgo treatment in patients with cerebral insufficiency and related disorders, came up with positive results, which contrast sharply with our own results. This inconsistency prompts us to search for satisfactory explanations. To what extent do the effects found in other trials differ from the effects that we have reported? And can any inconsistencies in the results be attributed to differences concerning the study design and conduct? First we will attempt to contrast several relevant aspects of the study design, and subsequently we will compare the reported effect sizes.

4.1 STUDY POPULATION

First of all it should be mentioned that the majority of the earlier trials failed to make a clear demarcation between various diagnostic categories nominated for admission into the study (nature of the phenomenon, degree of severity). Many studies recruited patients who could be caught under the broad concept of 'cerebral insufficiency', which includes also non-cognitive phenomena (e.g., tinnitus). Only some of the more recent trials focused on dementia (or even more specifically delineated types of dementia, such as Alzheimer's disease, vascular dementia), and/or memory impairment, following the currently accepted diagnostic standards. Generally speaking, however, the later studies showed positive effects of Ginkgo as well.

In most of the previous trials out-patients were recruited. This may have hampered the maintenance of an adequate level of compliance. Some studies were based on hospitalized patients. Recruitment from homes for the elderly or nursing homes was reported

only occasionally. Whereas in our trial the mean age of the participants exceeded 80 year, the preponderance in most of the trials was on patients aged 60-80, or even younger (especially for the studies focussing on non-cognitive signs and symptoms of cerebral insufficiency). Only a few French studies focussed on the same high age group.

A remarkable difference between our trial and some recent other trials regards the severity of the cognitive dysfunction of the study participants. The patients recruited by Kanowski et al, for instance, were covered by a more favourable range of eligible baseline scores for SKT and MMSE than our study subjects, whereas all these patients were classified as dementia cases. Usually a MMSE-score of 22 or 23 is regarded a sensible cut-off point between mild cognitive dysfunction and a more serious disorder, indicative for dementia. If we had applied the MMSE as the only screening tool to identify people with dementia, 76% of the current AAMI-group would have been identified as cases of dementia, given a MMSE baseline score ≤ 22 . We believe that we indeed have underdiagnosed dementia in our trial. The difficulties experienced with the application in our research setting of the DSM-III-R criterion C for dementia, are probably the main reason for this underdiagnosis. This criterion, which refers to the degree of interference of cognitive disturbances with work, social activities and functions, and ADL-functions, has been decisive for the classification as either dementia or AAMI case. We cannot imagine, however, that these diagnostic complications have seriously affected the validity of our trial results. Recently, a secondary analysis of the data of two previous Ginkgo trials^{6,19} revealed an equal level of efficacy in Ginkgo in patients with a mild and with a serious degree of cognitive impairment, according to MMSE baseline score²⁰.

4.2 SAMPLE SIZE AND STATISTICAL POWER

The sample sizes of the previous Ginkgo trials varied from 20 to 236 (number of subjects analyzed). Our trial harboured 214 participants (48 placebo and 166 Ginkgo), but only 123 were available for the main analysis after 24 weeks of intervention (79 Ginkgo and 44 placebo). Of the 7 previous trials that were based on a larger sample size, 2 could be characterized as dementia trials^{6,21}. We identified 3 trials with more than 200 subjects in the analysis, 4 trials with 150-200 analyzed participants, and 1 trial with 100-150 randomized patients. All other trials included less than 100 participants. From the perspective of statistical power, future trials should preferably contain at least 100 subjects in each treatment arm.

4.3 INTERVENTION

The study medication used in the Maastricht Ginkgo Trial was not much different from that in earlier trials. Almost all published Ginkgo trials have investigated the effects of the special extract EGb 761, that was also used in our trial, or LI 1370, with a slightly different concentration of Ginkgo flavonol glycosides. We copied the most common mode of administration of the study medication reported from previous trials (oral, tablets). The duration of the intervention and the follow-up of the participants was longer in our trial than in most previous trials. Only a few trials reported a similar (6 months)²²⁻²⁶, or even a longer (12 months)^{21,27,28} length. Although Ginkgo was contrasted with placebo in most other trials, the efforts needed to compose a placebo that really mimics the verum preparation, may have been underestimated. Therefore, we can imagine that the actual intervention contrast and, as a consequence, the blinded nature of the outcome measurements have been compromised in a lot of tri-

als. Moreover, we have noticed that it is not yet common practice for Ginkgo trials to check and report the achieved level of blinding. Although we find it hard to prove this, we assume that the compliance with the intervention regimen has also been better in our trial than in most other trials, in particular trials that made use of non-institutionalized populations. We noticed the scarce information on therapy compliance as another neglect of most Ginkgo trial reports. If we correctly assume that many previous trials failed to achieve the level of compliance realized in our trial, this means that they have missed the opportunity to show even larger effect sizes than actually were reported.

4.4 OUTCOME MEASUREMENTS

Only a limited number of the previous Ginkgo trials measured outcomes at all three levels that are prescribed by the current guidelines (neuropsychological performance, clinical symptoms, behaviour), and included a responder analysis based on a combination of these three levels of assessment. With regard to the individual outcome measures, it should be noted that most of the primary outcome measures that we used in our trial, have been applied regularly in Ginkgo trials previously; the Syndrom Kurz Test, for instance, was used in 5 trials, 4 of them dementia trials, the clinical global impression (CGI) in 4 trials, and the Nürnberger-Alltags-Aktivitäten-Skala (NAI-NAA) in 2 trials. The same holds for the secondary outcome measures, e.g., the Sandoz Clinical Assessment-Geriatric scale (SCAG; 5 trials), the trail making test (NAI-ZVT or a comparable number-connection test; 7 trials), the digit span test (NAI-ZN or a comparable format; 5 trials), the geriatric depression scale (1 trial). Unfortunately, most of the available trial reports keep us uninformed about the actual measurement conditions, which inevitably will have influenced the

measurement results and trial outcomes. We already questioned the blinded nature of the outcome measurements in many trials.

4.5 EFFECT SIZES

The effect sizes found in previous trials have been presented in chapter 2. Whereas in our trial the effect size for the SKT (difference in change over time between Ginkgo and placebo groups; scale: $-27 - +27$) amounted to $+0.4$ in favour of the Ginkgo group, the effect sizes reported in earlier trials ranged from $+1.4$ to $+7.0$ ^{22,29-32}. All these effects appeared to be statistically significant. Two trials showed an effect-size that outranged the level that was defined as a clinical meaningful change at the individual level in our trial (≥ 3.0)^{29,30}. In one of the best studies, published by Kanowski et al, 38% vs 18% of the Ginkgo and placebo groups, respectively, fulfilled the response criterion defined for that particular trial (improvement ≥ 4.0 points)²². In our trial 19% vs 11% responded according to the same criterion.

Based on the clinical global impression of change (CGI; scale: $1 - 7$), two trials reported a statistically significant effect in favour of the Ginkgo group, viz. improvement (score ≤ 3) in 85% vs 35%¹⁹, and even much improvement (score ≤ 2) in 32% vs 17%⁶ of the Ginkgo and the placebo groups, respectively. A borderline effect was found in a small-sized trial³⁰, and no effect at all in another trial²¹. In our trial 25% vs 23% improved according to the CGI, and we found mean scores of 4.3 vs 4.3.

Haase et al reported a relatively large effect of 4 weeks of intravenous Ginkgo administration on the level of daily activities, measured by means of the NAI-NAA (scale: $-40 - +40$): $+7.5$ in favour of the Ginkgo group¹⁹. No effect was found in another trial: 35% vs 38% of the participants showed any improvement³³. In our trial the effect size after 24 weeks of therapy was -0.0 for the NAA

scored by the nursing staff, and +1.3 points for the NAA scored by the participants themselves. 37% Vs 32% of the study subjects showed any improvement according to the nursing staff, and 35% vs 25% according to the patients themselves.

The effect sizes found in previous trials for the SCAG (mean value; scale: $-6.0 - +6.0$) varied from +0.3 to +1.6 in favour of the Ginkgo group^{29,32,34,35}, whereas one trial reported improvement for 29% vs 13% of the study participants³⁶. We found an effect size of +0.04 and any improvement in 49% vs 43% of the study participants.

A statistically significant effect in favour of Ginkgo with any form of a number-connection test was reported in 3 trials³⁷⁻³⁹, with effect sizes varying from +8 to +13 seconds. Four other trials reported a positive trend, but not a statistically significant effect for the NAI-ZVT^{30,31,33,40}. We registered an effect of +18 seconds, adjusted for confounding, which was however not statistically significant. Digit span testing gave a statistically significant positive result in 2 trials, a positive trend in two other trials, and a statistically significant effect in favour of placebo in one trial. We found no difference between Ginkgo and placebo.

The effect of Ginkgo on depressive mood was evaluated by means of the Geriatric Depression Scale (GDS) only once, but the results were reported⁴¹. In four trials quite impressive, statistically significant effects were measured by means of the Hamilton Depression Scale (HAM-D) and Zung's Self-Rating Depression Scale (SDS). With the GDS, we found a non-significant effect of +0.5 (scale: $-15 - +15$).

While reviewing the effect sizes, we noticed that many trials showed a degree of improvement (percentage improved, positive change of mean score over time) which exceeded the one that we recorded in our trial, not only for the Ginkgo users but also for the placebo users. Although also in our trial part

of the individual participants showed improvement, we did not find the overall placebo effect that was reported in other trials. We have no satisfactory explanation for this difference in study outcomes.

5 PERSONAL VIEW

Our attempts to reconcile the results of our trial with those of previously published Ginkgo trials have been quite unsuccessful. Nevertheless, in order to formulate a summary judgement on the efficacy of Ginkgo in patients with dementia and related disorders, we have to integrate the evidence based on our own trial with that available from other clinical trials. Research that focuses on the cellular, metabolic and physiological effects of Ginkgo compounds is another important source of evidence, primarily for the plausibility of clinical effects, and not for the clinical efficacy as such. Many studies into the bioavailability, the toxicology, and the experimental and clinical pharmacology of Ginkgo extracts have been published and reviewed in the past decades⁴²⁻⁴⁹. In general, these studies showed very promising results. The fact that we lack the scientific background and expertise that is needed for a critical assessment of these results, has been a major reason for not taking this 'preclinical' evidence into consideration. Our belief in the efficacy in Ginkgo before the start of our trial was quite high, high enough at least to decide to direct part of the investments in our own trial towards the measurement and analysis of the dose-dependency and the persistency of the presumed effect. By the way, we have never seriously tried to quantify our prior belief in Ginkgo, which was mainly based on what we learned from the criteria-based systematic review, which we referred to previously⁸. If we had done so, the prior probability of a beneficial Ginkgo effect, would presumably have varied between

30% and 70%, dependent on the level of in-born scepticism of each of the team members involved. Since the start of the Maastricht Ginkgo Trial the results of some 10 new trials have been published. The methodological quality of the majority of these trials was considerably higher than that of the older ones. Therefore, they harbour a disproportionate high amount of evidence. In spite of the somewhat smaller effects sizes reported for most of these trials, we feel that they have raised our belief in the efficacy of Ginkgo. In a recent POEM-statement (Patient-Oriented Evidence that Matters) in the Journal of Family Practice, that was largely inspired by the publication of the trial of Le Bars et al²¹, Stevermer and Lindbloom made the following recommendations for clinical practice⁵⁰: *"This particular extract of Ginkgo biloba, EGb 761, appears to have a modest stabilizing effect in the general functional decline of otherwise healthy patients with dementia. EGb appears to be as safe as placebo, although the small number of patients and short period limits the ability to detect uncommon events. The changes reported are of a similar magnitude to those seen with tacrine and donepezil, two currently available prescription medications, that, locally, cost three times more than Ginkgo extracts. Whether Ginkgo is safer or more effective than these medications is not clear. Recommendations to patients should be made with caution, since Ginkgo biloba does not face the regular scrutiny of prescription medications. Nonetheless, it appears that EGb may have some beneficial effects in demented individuals". Maybe this comment was a little bit too reticent, in particular with regard to the tolerability of Ginkgo, although serious concerns about its safety are published occasionally*⁵¹.

We think that the - negative - results of our own trial give additional weight to the already existing evidence, in particular in view of the following methodological aspects: the ran-

dom character of the study treatment allocation, the control exerted by the placebo preparation used, the blinded outcome assessment, and the low dropout rate. We have expressed some doubts with regard to the choice of our study population and the adequacy of some outcome measurements in our trial. On the other hand, the remarkable consistency of the negative outcomes, based on a large variety of data collection methods, strengthens our belief that exposure to Ginkgo indeed failed to evoke an observable response in our study population. All together, we conclude that the results of our trial have scaled down our belief in the efficacy of Ginkgo. Again, we are hesitating to quantify our posterior belief. We realize that our negative results cannot fully neutralize the positive results of previous trials. For this moment we find it hard to discredit the outcomes of some of the most recent trials in particular. These trials had many design aspects in common with our study, although, obviously, their reports hide more secrets for us than our own trial. As yet, we cannot refute the possibility that our trial has been an 'outlier by chance'. Therefore, at the risk of sounding somewhat cheap and standard, we have to conclude with the plea for one or more new, rigorously designed trials, to be planned and conducted under strict scrutiny, and which should include at least several hundreds of patients.

REFERENCES

1. Warburton DM. Psycho-pharmacologie clinique de l'extrait de Ginkgo biloba. *Presse Med* 1986; 15: 1595-604.
2. Weiß H, Kallischnigg G. Ginkgo-biloba-Extrakt (EGb 761). Meta-Analyse von Studien zum Nachweis der therapeutischen Wirksamkeit bei Hirnleistungsstörungen bzw. arterieller Verschlusskrankheit. *Münch Med Wochenschr* 1991; 133: 138-42.

3. Herrschaft H. Zur klinischen Anwendung von *Ginkgo biloba* bei dementiellen Syndromen (Hirnleistungsstörungen bei vasculärer oder degenerativer ZNS Erkrankung). *Pharm Unserer Zeit* 1992; 21: 266-75.
4. Schulz V, Hänsel R. *Rationale Phytotherapie. Ratgeber für die ärztliche Praxis*. Ch.2: Zentrales Nervensystem. § 2.1. Ginkgo bei Hirnleistungsstörungen. Berlin: Springer, 1996.
5. Letzel H, Haan J, Feil WB. Nootropics: efficacy and tolerability of products from three active substance classes. *J Drug Dev Clin Pract* 1996; 8: 77-94.
6. Kanowski S. Ginkgo-biloba-Spezialextrakt. Nachgewiesene Wirksamkeit im Indikationsbereich Demenz. *Münch Med Wochenschr* 1997; 139: 39-42.
7. Schulz V, Hubner WD, Ploch M. Clinical trials with phyto-psychopharmacological agents. *Phytotherapy* 1997; 4: 379-87.
8. Kleijnen J, Knipschild P. *Ginkgo biloba* for cerebral insufficiency. *Br J Clin Pharmacol* 1992; 34: 352-8.
9. Hopfenmüller W. Nachweis der therapeutischen Wirksamkeit eines *Ginkgo biloba* Spezialextraktes. Meta-Analyse von 11 klinischen Studien bei Patienten mit Hirnleistungsstörungen im Alter. *Arzneimittelforschung* 1994; 44: 1005-13.
10. Pocock SJ. *Clinical trials. A practical approach*. Chichester: John Wiley & Sons, 1993.
11. Meinert CL. *Clinical trials: design, conduct and analysis*. New York: Oxford University Press, 1986.
12. Ter Riet G, Kleijnen J, Knipschild P. Acupuncture and chronic pain. A criteria-based meta-analysis. *J Clin Epid* 1990; 43: 1191-9.
13. Koes BW, Bouter LM, Van der Heijden GJMG, Knipschild PG. Physiotherapy exercises and back pain. *BMJ* 1991; 302: 1572-6.
14. Van der Heijden GJMG, Beurskens AJHM, Koes BW, Assendelft WJJ, De Vet HCW, Bouter LM. Traction for back and neck pain: a blinded review. *Physical Therapy* 1995; 75: 93-104.
15. De Vet HCW, De Bie RA, Van der Heijden GJMG, Verhagen AP, Sijpkens P, Knipschild PG. Systematic reviews on the basis of methodological criteria. *Physiotherapy* 1997; 83: 184-9.
16. Knipschild PG, Hörr R, Van Rossum E, Van Dongen MCJM. Optimization of placebos for double-blind clinical trials: experience with a phytopharmaceutical. *Arzneim.-Forsch./Drug Res.* 1998; 48: 1033-6.
17. Detsky AS, Naylor CD, Rourke K, McGeer AJ, L'Abbe KA. Incorporating variations in the quality of individual randomized controlled trials into meta-analysis. *J Clin Epid* 1992; 45: 255-65.
18. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; 315: 629-34.
19. Haase J, Halama P, Hörr R. Wirksamkeit kurzdauernder Infusionsbehandlungen mit Ginkgo-biloba-Spezialextrakt EGB 761 bei Demenz vom vaskulären und Alzheimer-Typ. *Z Gerontol Geriatr* 1996; 29: 302-9.
20. Hörr R, Kieser M. Ginkgo-biloba-Spezialextrakt EGB 761® - ein Antidementivum. Wirksam auch bei deutlicher Krankheitsausprägung. *Fortschr Med* 1998; 16: 39-40.
21. Le Bars PL, Katz MM, Berman N, et al. A placebo-controlled, double-blind, randomized trial of an extract of *Ginkgo biloba* for dementia. *JAMA* 1997; 278: 1327-32.
22. 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.
23. Chartres J-P, Bonnan P, Martin G. Réduction de posologie de médicaments psychotropes chez des personnes âgées vivant en institution. Étude à double-insu chez des patients prenant soit de l'extrait de *Ginkgo biloba* 761 soit du placebo. *Psychologie Médicale* 1987; 19: 1365-75.
24. Rai GS, Shovlin C, Wesnes KA. A double-blind, placebo controlled study of *Ginkgo biloba* extract ('tanakan') in elderly outpatients with mild to moderate memory impairment. *Curr Med Res Opin* 1991; 12: 350-5.

25. Augustin P. Le Tanakan en gériatrie. Étude clinique et psychométrique chez 189 malades d'hospice. *Psychologie Médicale* 1976; 8: 123-30.
26. Gräbel E. Einfluß von *Ginkgo biloba* Extrakt auf die geistige Leistungsfähigkeit. Doppelblindstudie unter computerisierten Meßbedingungen bei Patienten mit Zerebralsuffizienz. *Fortschr Med* 1992; 110: 73-6.
27. Franco L, Cuny G, Nancy FMC. Étude multicentrique de l'efficacité de l'extrait de *Ginkgo biloba* (EGb 761) dans le traitement des troubles mnésiques liés à l'âge. *La Revue de Gériatrie* 1991; 16: 191-5.
28. Taillandier J, Ammar A, Rabourdin JP, et al. Traitement des troubles du vieillissement cérébral par l'extrait de *Ginkgo biloba*. Étude longitudinale multicentrique à double insu face au placebo. *Presse Med* 1986; 15: 1583-7.
29. Hofferberth B. The efficacy of EGb 761 in patients with senile dementia of the Alzheimer type, a double-blind, placebo-controlled study on different levels of investigation. *Human Psychopharmacology* 1994; 9: 215-22.
30. Maurer K, Ihl R, Dierks T, Frölich L. Clinical efficacy of *Ginkgo biloba* special extract EGb 761 in dementia of the Alzheimer type. *J Psychiatr Res* 1997; 31: 645-55.
31. Halama P. *Ginkgo biloba*. Wirksamkeit eines Spezialextrakts bei Patienten mit zerebraler Insuffizienz. *Munch Med Wochenschr* 1991; 133: 190-4.
32. Halama P, Bartsch G, Meng G. Hirnleistungsstörungen vaskulärer Genese. Randomisierte Doppelblindstudie zur Wirksamkeit von *Ginkgo biloba* Extrakt. *Fortschr Med* 1988; 106: 408-12.
33. Oswald WD, Hörr R, Oswald B, Steger W, Sappa J. Zur Verbesserung fluider, kognitiver Leistungen mit *Ginkgo-biloba*-Spezialextrakt EGb 761® bei Patienten mit leichten bis mittelschweren Hirnleistungsstörungen im Alter. *Zeitschrift für Gerontopsychologie und -psychiatrie* 1997; 10: 133-46.
34. Taillandier J, Ammar A, Rabourdin JP, et al. Traitement des troubles du vieillissement cérébral par l'extrait de *Ginkgo biloba*. Étude longitudinale multicentrique à double insu face au placebo. *Presse Med* 1986; 15: 1583-7.
35. Weitbrecht WV, Jansen W. Doubleblind and comparative (*Ginkgo biloba* versus placebo) therapeutic study in geriatric patients with primary degenerative dementia - a preliminary evaluation. In: Agnoli J, Rapin R, Scapagnini V, Weitbrecht WV, eds. *Effects of Ginkgo Biloba Extract on organic cerebral impairment*. Montrouge: John Libbey Eurotext Ltd., 1985: 91-99.
36. Pidoux B, Bastien C, Niddam S. Clinical and quantitative EEG double-blind study of *Ginkgo biloba* extract (GBE). *J Cereb Blood Flow Metab* 1983; 3: S556 - S557.
37. Hofferberth B. Einfluß von *Ginkgo biloba* Extrakt auf neurophysiologische und psychometrische Meßergebnisse bei Patienten mit hirnganischem Psychosyndrom. Eine Doppelblindstudie gegen Plazebo. *Arzneimittelforschung* 1989; 39: 918-22.
38. Hofferberth B. *Ginkgo-biloba*-Spezialextrakt bei Patienten mit hirnganischem Psychosyndrom. Prüfung der Wirksamkeit mit neurophysiologischen und psychometrischen Methoden. *Munch Med Wochenschr* 1991; 133 (Suppl. 30): S30 - S33.
39. Brüchert E, Heinrich SE, Ruf-Kohler P. Wirksamkeit von LI 1370 bei älteren Patienten mit Hirnleistungsschwäche. Multizentrische Doppelblindstudie des Fachverbandes Deutscher Allgemeinärzte. *Munch Med Wochenschr* 1991; 133 (Suppl. 30): S9 - S14.
40. Hartmann A, Frick M. Wirkung eines *Ginkgo*-Spezial-Extraktes auf psychometrische Parameter bei Patienten mit vaskulär bedingter Demenz. *Munch Med Wochenschr* 1991; 133 (Suppl. 30): S23 - S25.
41. Israël L, Dell'Accio E, Martin G, Hugonot R. Extrait de *Ginkgo biloba* et exercices d'entraînement de la mémoire. Evaluation comparative chez des personnes âgées ambulatoires. *Psychologie Médicale* 1987; 19: 1431-9.
42. Christen Y, Costentin J, Lacour M, eds. *Effects of Ginkgo biloba Extract (EGb 761) on the central nervous system*. Paris: Elsevier, 1992.
43. Christen Y, Droy-Lefaix M, Macias-Nunez J, eds. *Effects of Ginkgo biloba Extract (EGb 761) on neuronal plasticity*. Paris: Elsevier, 1996.

44. DeFeudis FV. *Ginkgo biloba* extract (EGb 761). *From chemistry to the clinic*. Wiesbaden: Ullstein Medical, 1998.
45. Ferradini C, Droy-Lefaix M, Christen Y, eds. *Ginkgo biloba Extract (EGb 761) as a free-radical scavenger*. Paris: Elsevier, 1993.
46. Clostre F, DeFeudis F, eds. *Cardiovascular effects of Ginkgo biloba Extract (EGb 761)*. Paris: Elsevier, 1994.
47. Hori T, Ridge RW, Tulecke W, Del Tredici P, Trémouillaux-Guiller J, Tobe H, eds. *Ginkgo biloba. A global treasure. From biology to medicine*. Tokyo: Springer / The Botanical Society of Japan, 1997.
48. Packer L, Christen Y, eds. *Ginkgo biloba extract (EGb 761): lessons from cell biology*. Paris: Elsevier, 1998.
49. Anonymous. *Tebonin[®] (EGb 761). Ginkgo biloba Special Extract*. The result of 25 years of research. Karlsruhe: Dr. Willmar Schwabe GmbH & Co, 1991.
50. Stevermer JJ, Lindbloom EJ. *Ginkgo biloba* for dementia. *J Fam Pract* 1998; 46: 20.
51. Vale S. Subarachnoid haemorrhage associated with *Ginkgo biloba*. *Lancet* 1998; 352: 36.

SUMMARY

Chapter 1 provides an introduction to this thesis. It highlights the considerations and motives that led to the decision to initiate the Maastricht Ginkgo Trial. A description of the design and the results of this trial takes up the major part of the current thesis. Some background information is given on the phytopharmaceutical use of *Ginkgo biloba* products in historical perspective, the pharmacological mechanisms that are supposed to underlie the clinical efficacy of *Ginkgo biloba* extracts ('Ginkgo'), the introduction of various commercially traded *Ginkgo biloba* special extracts in western countries and the development of their role on the health market in modern times. Experiences obtained from previous Ginkgo trials in patients with cerebral insufficiency, including dementia, are identified as the main rationale of the architecture of the Maastricht Ginkgo Trial.

Chapter 2 contains a review of the clinical efficacy of Ginkgo in patients with manifestations of cerebral insufficiency. An extensive search of the relevant literature yielded 55 randomized clinical trials that we considered eligible for inclusion in the review. In summarizing the evidence contained by these trials, we decided not to stick to the principles of criteria-based systematic reviewing and statistical pooling of study outcomes. Several considerations made us to abandon these generally accepted methods, such as the lack of completeness and transparency of many study reports, and the uncertainties that surround the diagnosis and outcome measurement of dementia and related disor-

ders; these, in turn, have given rise to quite unsettled diagnostic and therapeutic evaluation guidelines as well as the actual application of a very heterogeneous collection of patient definitions and outcome measures in clinical trials. The approach chosen for our review implied the categorization of all identified trials according to the main health target of interest, followed by a detailed description of the design characteristics and the results of each trial, guided by a fixed format. Of the 55 trials, 9 focussed on dementia, 8 on serious cognitive impairment, 12 on non-cognitive signs and symptoms of cerebral insufficiency, 5 on subjective cognitive and non-cognitive symptoms of cerebral insufficiency, 3 on depression, and 18 on unspecified forms of cerebral insufficiency. Beneficial effects of Ginkgo were reported for almost all trials in each of the categories, although not invariably for all outcome measures. In general, older trials showed larger effects than more recent ones, whereas the newer trials tended to be superior with regard to methodological quality. We conclude that the available evidence strongly suggests that Ginkgo is efficacious in patients with dementia and other types of cerebral insufficiency, although only a limited number of trials fulfils the current standards for evaluation and reporting. Moreover, the clinical meaningfulness of the reported effect sizes remains difficult to interpret.

Chapter 3 presents the design of the Maastricht Ginkgo Trial. The aim of this intervention study was to assess the efficacy of vary-

ing amounts of therapeutic dosage and varying duration of intervention with orally administered *Ginkgo biloba* extract (EGb 761) in patients with mild to moderate dementia or age-associated memory impairment. The study was designed as a randomized, placebo-controlled, double-blind, parallel trial.

The study participants were recruited from homes for the elderly in the southern part of the Netherlands, by means of a multi-stage screening procedure. Enrolment in the trial required the presence of either dementia (Alzheimer dementia (AD), vascular dementia (VD), or a mixed type of dementia) or age-associated memory impairment (AAMI). The diagnosis of dementia was made by means of the SIDAM, a structured screening interview which also involves assessment by informants (i.e. nursing staff), and which fully covers the DSM-III-R and ICD-10 criteria for dementia. In addition to the SIDAM we used the Syndrom Kurz Test (SKT), a short test battery on memory and attention, to determine the severity of dementia (scores: 8-23). AAMI-patients were identified among non-demented subjects by means of objective psychometric testing (SKT, scores 8-23), and subjective assessment of the impact of memory problems in daily life (modified MAC-Q questionnaire, score ≥ 11). To be eligible, the patients of both diagnostic subtypes had to fulfil various additional entry criteria, e.g., with regard to age (≥ 50 years), written informed consent, absence of severe depression (GDS < 11), absence of serious co-morbidity and relevant co-interventions, and absence of placebo-reponse and non-compliance during the run-in period that preceded the trial. Eventually, 214 eligible subjects (63 with dementia and 151 with AAMI) were recruited in 39 homes for the elderly. The study treatment consisted of the oral administration of film-coated tablets, that contained either 0 mg (placebo), or 80 mg, or 120 mg of the *Ginkgo biloba* extract EGb

761. The development of a high-quality placebo received special attention. Each day two tablets were administered (morning and evening, during the meal), as a rule under close supervision of a nursing staff member. Packaging and labelling of the study medication were performed in accordance with the prevailing guidelines and requirements.

Allocation to the treatment alternatives (Ginkgo 240 mg, Ginkgo 160 mg, placebo) was guided by a two-stage concealed randomization procedure. The first randomization step immediately preceded the baseline measurements, and was conducted according to a 2 : 2 : 1 allocation ratio, within predefined strata for diagnostic subtype (dementia vs AAMI) and estimated change in memory during the 6-month pretreatment phase (improvement or no change vs impairment). Only participants who completed 12 weeks of Ginkgo treatment were involved in the second randomization step, that was required to enable the analysis of the durability of any intervention effect. These subjects were allocated to continued treatment with Ginkgo (same dose) or placebo for the remaining 12 weeks of the trial.

In order to enable adjustment for confounding in the phase of data analysis, several factors that might predict the prognosis of dementia and cognitive impairment, such as age, gender, and the history of some relevant biological life events (BLEs), were measured at baseline.

To evaluate the intended effects of Ginkgo primary outcome measures were defined at three different levels of performance and functioning, in accordance with recent standards for the evaluation of anti-dementia drugs and cognition enhancers. In addition to these primary outcome measures several additional outcome measures were chosen, which we regarded of secondary importance as far as the decision on the efficacy of Ginkgo is concerned. The primary outcome measures were: the Syndrom Kurz

Test (SKT), a short test battery of memory and attention, at the level of neuropsychological functioning; the Clinical Global Impression of change in cognitive functioning (CGI-2), at the level of clinical presentation; and the Nürnberger Alters Inventar Alltags-Aktivitäten Skala (NAI-NAA, assessment by nursing staff), a rating scale addressing various cognition-related activities of daily life, at the level of behavioural functioning. The secondary outcome measures included several psychometric tests (trail making test (NAI-ZVT-G), digit span test (NAI-ZN-G), word list reproduction and recognition (NAI-WL)), clinical assessments (geriatric symptoms (SCAG), depressive mood (GDS, 15-item version), self-perceived health status (report mark), and self-perceived memory status (report mark)), and behavioural assessment (activities of daily life (NAI-NAA, self-assessment)). Baseline data for all outcome parameters were collected before the start of the intervention. Outcomes were assessed after 12 and 24 weeks of treatment for all outcome parameters, and at 4, 8, and 18 weeks of follow-up for a subcategory of outcome parameters (SKT, CGI, SCAG, health status, memory status). In addition to the intended outcome phenomena, we assessed several other phenomena of interest at each follow-up measurement time-point (4, 8, 12, 18, and 24 weeks), viz. the occurrence of psychological life events (PLE), that might influence the cognitive status, the compliance with study medication (pill count), and the occurrence of (serious) adverse events (checklist of common health complaints; spontaneous reporting of serious complications by the nursing staff). Manifestations of addiction (symptom checklist) were assessed after 18 and 24 weeks of treatment, and exposure to co-interventions (co-medication, change in care) after 12 and 24 weeks. Finally, the thoroughness of blinding was checked after 4 and 18 weeks. At each measurement time point we received information with regard to

intended outcomes and additional phenomena through two different sources: interviews with the study participants, conducted by a team of trained interviewers; and observations by members of the nursing staff in each of the participating homes for the elderly.

The confirmatory analysis of the study outcomes consisted of an intention-to-treat analysis of the 24-week follow-up data regarding the primary outcome measures, for those participants who received Ginkgo or placebo during the entire intervention period. An overall OLS-test preceded tests for each of the separate outcome parameters. Moreover, we performed a responder analysis, focussing on study subjects with clinically meaningful effects for at least two of the three primary outcome measures. In addition, descriptive analyses were conducted to study the effects for the secondary outcome parameters, the effects after 12 weeks of treatment, the dose-effect relationship, the effect of 24-week versus 12-week treatment, the effects in various subgroups, and the effects for valid cases only (per-protocol analysis).

Chapter 4 describes the main results of the Maastricht Ginkgo Trial. These results concerned the primary outcome measures that we used to evaluate the efficacy and safety of the *Ginkgo biloba* extract EGb 761. Within 39 homes for the elderly, the multi-stage screening process yielded 214 study subjects who fulfilled all enrolment criteria and who eventually could be randomized. The mean age of the study population was 82.6 years; 180 were female, and 34 male. 63 Patients were diagnosed with dementia, and 151 with age-associated memory impairment (AAMI). Eighteen participants withdrew prematurely from the study population, 13 during the first, and again 5 during the second 12-week intervention period, most of them due to death or severe co-morbidity. The dropout rates were 11%, 6% and 8% for

the initial Ginkgo 240 mg, Ginkgo 160 mg, and placebo groups, respectively. During the entire study period we registered 28 non-compliers (including the dropouts). These were also quite equally distributed over the treatment groups. We found no relevant baseline differences between the distinct treatment groups. A check for blinding revealed no clear association between the presumed and the actual study medication, for any of the categories of assessors involved in the trial.

123 Of the 214 randomized subjects were eligible for the confirmatory analysis of the trial results, since they received either Ginkgo ($n=79$; 240 and 160 mg/day combined) or placebo ($n=44$) during the entire 24-week intervention period. An overall-test, which covered all three primary outcome measures, indicated no effect of Ginkgo ($T_{OLS} = 0.34$; $p=0.37$). For each of the separate outcome measures no effect was found either. Memory and attention measured by means of the Syndrom Kurz Test (SKT, scale of status score: 0-27; underlining represents the most favourable extreme; deteriorated slightly in both the combined Ginkgo and the placebo group: -0.82 vs. -1.23 , resulting in a for confounding adjusted effect of $+0.41$ (90% CI: -0.86 ; $+1.68$), in favour of Ginkgo. The same happened with the level of cognition-related activities of daily life, measured by the Nürnberger Alters Inventar Alltags-Aktivitäten Skala (NAI-NAA; scale: 20-60): -1.37 vs. -1.36 , with an adjusted difference of -0.36 (90% CI: -1.94 vs. $+1.22$), in favour of the placebo group. The Clinical Global Impression of change in cognitive function (CGI; 1-7, scored by the nursing staff, also did not show an effect: 4.29 vs. 4.34 , giving an unadjusted effect of $+0.05$ (90% CI: -0.31 ; $+0.41$). Despite of a larger statistical power (166 subjects in the Ginkgo group, and 48 in the placebo group) no effects were found after 12 weeks of intervention either: SKT: $+1.00$ (90% CI: -0.05 ; $+2.05$); NAI-

NAA: $+0.25$ (90% CI: -0.98 ; $+1.48$); GCI: -0.03 (90% CI: -0.27 ; $+0.21$). After 24 weeks of treatment 14% and 18% of the subjects in the combined Ginkgo and placebo groups, respectively, showed a clinical relevant response to the study treatment (Chi-square = 0.39 ; $p = 0.53$). Neither the dementia subgroup, nor the AAMI-subgroup, nor any other subgroup evaluated experienced a significant effect of Ginkgo. There was no dose-effect relationship, and also no effect of prolonged Ginkgo treatment. Ginkgo treatment was also not associated with the occurrence of (serious) adverse events. A critical assessment of the adequacy and comparability of study populations, interventions and outcomes did not reveal satisfactory explanations for the absence of a treatment effect.

Chapter 5 describes the results of the Maasricht Ginkgo Trial with regard to a series of additional outcome measures, that included psychometric tests, clinical assessment, and behavioural assessment. After 24 weeks of continuous treatment with either Ginkgo or placebo no statistical significant and meaningful differences in the change score for any of the outcome parameters could be detected. After adjustment for baseline differences we found the following effects, with a positive effect indicating that the Ginkgo group ($n=79$) performed better, and a negative sign that the placebo group ($n=44$) performed better: trail making test (NAI-ZVT-G; scale of status score: 0-300 seconds): $+18.4$ (90% CI: -0.7 ; $+37.5$); digit span test (NAI-ZN-G; scale: 0-17): -0.01 (90% CI: -0.54 ; $+0.52$); word list (NAI-WL; scale: 0-16): $+0.12$ (90% CI: -0.69 ; $+0.93$); geriatric symptoms (SCAG; scale: 10-7.0): $+0.01$ (90% CI: -0.22 ; $+0.24$); depressive mood (GDS; scale: 0-15): $+0.54$ (90% CI: -0.46 ; $+1.54$); health status (report mark; scale: 0-10): -0.15 (90% CI: -0.56 ; $+0.26$); memory status (report mark; scale: 0-10): -0.15 (90%

CI: -0.51 ; $+0.21$); and activities of daily life (NAI-NAA, self-assessment; scale: 20 -60): $+1.32$ (90% CI : $+0.13$; $+2.77$). At halftime, after 12 weeks of treatment, the subjects in the Ginkgo group (240 and 160 mg per day combined; $n=166$) performed slightly better than the subjects in the placebo group ($n=48$) with regard to self-reported activities of daily life (NAI-NAA; effect size: $+1.35$ (90% CI: $+0.22$; $+2.49$), but slightly worse with regard to self-perceived health status (effect size: -0.38 (90% CI: -0.73 ; -0.04)). The subjects who received the highest Ginkgo dose (240 mg/day) performed not better than the subjects who received the usual dose (160 mg per day). No beneficial effects of a prolonged duration of Ginkgo treatment - 24 weeks in stead of 12 weeks - were found either. Moreover, we could not detect any subgroup that benefitted from Ginkgo. Compared with the intention-to-treat analysis, a per-protocol analysis gave quite similar results. The findings for the secondary outcome measures appeared to be compatible with the findings regarding the primary outcome measures, which we presented in chapter 4. Based on these results we had to conclude that the *Ginkgo biloba* special extract EGb 761 does not seem to be effective in the treatment of elderly people with mild to moderate dementia or age-associated memory impairment.

Chapter 6 contains a general discussion on the findings presented in the previous chapters. With regard to the clinical trials on *Ginkgo biloba* and cerebral insufficiency that have been published before and after the start of the Maastricht Ginkgo Trial, we conclude that the existing evidence for a benefi-

cial effect of Ginkgo is quite substantial. Nevertheless, the majority of the evidence seems to originate from a minority of the trials, that focussed on dementia and appear to satisfy the latest requirements for the evaluation of nootropics and anti-dementia drugs. All these studies have been published quite recently, and show rather modest effects of Ginkgo. We cannot bring forward convincing proof for the possibility that selective publication has caused biased review results. At the other hand, this possibility cannot be ruled out at all, since most of the activities in the context of Ginkgo research have been carried out without strict scrutiny and supervision by an independent office. As far as the validity of the Maastricht Ginkgo Trial is concerned, we feel that the randomization procedure, the high-quality placebo intervention, the high level of compliance with the study medication, and the blinded nature of the outcome measures have all contributed to the evidential power of the study. We are inclined to express some hesitations with regard to the appropriateness of outcome measurement (choice of measures, conduct of measurements) and patient selection. However, we can hardly imagine that any flaws in this respect have caused the complete masking of a really existing Ginkgo effect.

Although we point to some dissimilarities in study design, especially with regard to the composition of the patient population, by the end we fail to reconcile our findings with the available evidence from earlier trials. One or more large-scaled, rigorously designed trials are needed to settle the dispute on the efficacy of Ginkgo in dementia and cognitive decline.

SAMENVATTING

Centraal in dit proefschrift staat de Maastrichtse Ginkgo Trial. Dit is een inmiddels afgesloten interventie-onderzoek naar de effectiviteit van een speciaal extract van de *Ginkgo biloba* boom (EGb 761) bij oudere mensen met beginnende dementie dan wel relatief ernstige geheugenstoornissen. Het onderzoek werd uitgevoerd door medewerkers van de vakgroep Epidemiologie van de Universiteit Maastricht, met financiële steun van het farmaceutisch bedrijf Willmar Schwabe Arzneimittel GmbH, in Karlsruhe, Duitsland.

De Ginkgo-boom heeft haar wortels in het Verre Oosten (China, Japan, etc.). Karakteristiek zijn met vooral de twee-lobbige bladeren - vandaar: 'biloba' - die door hun opvallende vorm allerlei associaties oproepen. Ze hebben bijvoorbeeld veel weg van olifantsoeren. Uit overlevering weten we dat extracten van de boom reeds vele eeuwen geleden werden gebruikt als phytotherapeuticum tegen tal van kwalen. In de 18e eeuw werd de Ginkgo-boom (opnieuw) geïntroduceerd in West-Europa. Vandaag de dag komt men hem overal als sierheester tegen. In 1965 werd in Duitsland begonnen met de productie van een hoogwaardig extract uit de bladeren van de Ginkgo-boom, op basis van een gestandaardiseerd proces van zuivering en concentratie. In dit speciale extract komen bepaalde klassen van verbindingen (Ginkgo flavonol glycosiden, terpeen trilactonen) in vaste hoeveelheden en verhoudingen voor. Onderzoek heeft sindsdien aangetoond dat de stoffen in dit extract - kortweg: 'Ginkgo' -

in potentie op verschillende manieren een gunstige werking op het lichaam kunnen hebben: verbetering van de bloedcirculatie via vaatverwijding, vermindering van de stropigheid van het bloed, en het tegengaan van het samenklonteren van bloedplaatjes; een gunstig effect op de celstofwisseling in het centraal zenuwstelsel, met als gevolg een verhoogde tolerantie tegen zuurstofschuld; beïnvloeding van de het neurotransmitter systeem in de hersenen; en preventie van de beschadiging van celmembranen door vrije radicalen. Op grond van de veronderstelde werking wordt geclaimd dat Ginkgo vooral heilzaam is bij zogenaamde 'doorbloedingsstoornissen' op verschillende plaatsen in het lichaam, met name in de benen (perifeer vaatlijden, 'etalage-been') en in de hersenen (cerebrale insufficiëntie: een verminderde hersenfunctie die zich uit in klachten als concentratie- en geheugenverlies, lusteloosheid, depressie, angst, duizeligheid, oorsuizen, en hoofdpijn). Preparaten op basis van Ginkgo-extract zijn inmiddels in diverse Europese landen officieel geregistreerd als geneesmiddel. In een land als Duitsland behoren ze al jarenlang tot de meest frequent voorgeschreven en verkochte geneesmiddelen. In 1994 vond in Duitsland goedkeuring als middel tegen dementie plaats. In Nederland is Ginkgo veel minder bekend. Toch treft men ook bij ons in de apotheek, bij de drogist, in de reformwinkel, en tegenwoordig ook in zogenaamde 'smartshops' allerlei Ginkgo-preparaten en producten met Ginkgo-bestanddelen aan. Onlangs werd in Nederland het eerste

Ginkgo-handelspreparaat als geneesmiddel (tegen perifere vaatlijden) geregistreerd.

De directe aanleiding voor het Maastrichtse onderzoek naar de effectiviteit van Ginkgo bij dementie en geheugenverlies was de volgende. Uit een systematisch literatuuronderzoek dat in 1992 door Kleijnen en Knipschild werd gepubliceerd, bleek dat op dat moment al heel wat gecontroleerde experimenten naar de effectiviteit van Ginkgo biloba waren uitgevoerd en gerapporteerd. Zo resulteerde een grondige speurtocht in 40 gerandomiseerde trials naar Ginkgo biloba bij cerebrale insufficiëntie op 39. Hiervan wezen op een positief effect van Ginkgo. Wat dat betreft onderscheidde Ginkgo zich in gunstige zin temidden van een hele reeks (andere) voedingssupplementen en alternatieve therapieën die op soortgelijke wijze tegen het licht werden gehouden. Op veel van de geïdentificeerde studies viel echter het nodige aan te merken. Daarom leek de tijd rijp voor een nieuwe studie, nu eens niet in Frankrijk en Duitsland, en uitgaande van een scherper gedefinieerde onderzoekspopulatie. De hoopgevende resultaten van eerdere studies rechtvaardigden bovendien de beslissing om aansluitend op de hoofdvraag van het nieuwe experiment - is ginkgo effectief bij beginnende dementie en geheugenklachten? - twee extra vragen te formuleren. Even aannemende dat ginkgo effectief is, gaat een hogere dosis dan gebruikelijk is gepaard met een sneller of sterker effect? En, opnieuw aannemende dat ginkgo effectief is, volstaat in dat geval een relatief korte kuur, of ebt het effect na het stoppen met de behandeling weer snel weg?

Tegen deze achtergrond werd de Maastrichtse Ginkgo Trial opgezet: een gerandomiseerd, dubbel-blind, placebo-gecontroleerd experiment. De deelnemers aan het onderzoek werden in de periode 1994-1996 geworven onder de bewoners van 39 verzor-

gingshuizen in Limburg en het oostelijk deel van Noord-Brabant. De selectie van geschikte onderzoekspersonen gebeurde via een aantal stappen. Centraal hierbij stond een screeningsonderzoek bij de kandidaat-deelnemers, bestaande uit een vraaggesprek en diverse neuropsychologische functietesten, alsmede een beoordeling door het verzorgend personeel in de instelling.

In **Hoofdstuk 1** worden de overwegingen en motieven belicht die aan de opzet en de uitvoering van de Maastrichtse Ginkgo Trial ten grondslag hebben gelegen. De historische betekenis van Ginkgo als phytotherapeuticum komt aan de orde. Voorts wordt aandacht besteed aan de werkingsmechanismen die mogelijk de basis vormen van het veronderstelde klinische effect van *Ginkgo biloba* extracten ('Ginkgo'). Ook wordt ingegaan op de wijze waarop diverse commercieel verhandelde Ginkgo-preparaten proberen zich een plaats te veroveren op de hedendaagse markt van gezondheidsproducten, met name in westerse landen. De ervaringen opgedaan met eerdere Ginkgo trials als gericht op patiënten met cerebrale insufficiëntie, worden genoemd als belangrijkste leidraad bij het opzetten van de Maastrichtse Ginkgo Trial.

Hoofdstuk 2 geeft een literatuuroverzicht van de klinische effecten van Ginkgo bij patiënten met verschijnselen van cerebrale insufficiëntie. Een grondige screening van de bestaande literatuur bracht 55 gepubliceerde trials aan het licht die voor opname in het overzicht in aanmerking kwamen. Bij de analyse van deze documenten werd noch voor de aanpak van systematische toetsing aan methodologische kwaliteitscriteria gekozen ('criteria-based systematic review', 'blinded review'), noch voor de aanpak die berust op statistische pooling van de uitkomsten gemeten in verschillende studies. Diverse overwegingen hebben ons doen afzien van

het toepassen van deze geaccepteerde methoden om de literatuur over interventiestudies samen te vatten. Het grootste struikelblok vormde het gebrek aan volledigheid en transparantie van veel onderzoeksrapporten, in combinatie met de onzekerheid over de vraag hoe dementie en verwante aandoeningen gediagnosticeerd moeten worden en hoe het beloop ervan geëvalueerd moet worden. Ook nu nog zijn de richtlijnen voor de diagnose van dementie en voor de evaluatie van interventies gericht op de bestrijding van dementie voortdurend in beweging. Dit uit zich onder meer in de grote verscheidenheid van patiënt-definities en uitkomstmaten die in de klinische trials op dit terrein gehanteerd worden. Voor de huidige review hebben we de volgende aanpak gekozen. We hebben alle relevante trials gegroepeerd op basis van de hoofd-indicatie voor behandeling met Ginkgo. Daarna hebben we per categorie de belangrijkste kenmerken van de opzet en van de uitkomsten van elke trial gedetailleerd in kaart gebracht, volgens een vast stramien. We slaagden erin 55 trials te identificeren. Hiervan richtten er zich 9 op dementie, 8 op een relatief ernstige mate van achteruitgang van de cognitieve functies, in het bijzonder de geheugenfunctie, 12 op niet-cognitieve symptomen van cerebrale insufficiëntie, 5 op een min of meer vaste combinatie van cognitieve en niet-cognitieve symptomen van cerebrale insufficiëntie, 3 op depressieve klachten, en 18 op niet nader gespecificeerde vormen van cerebrale insufficiëntie. Voor vrijwel alle trials binnen elk van de zojuist genoemde categorieën werd een positief effect gevonden, zij het niet steeds voor elk van de bestudeerde uitkomstmaten. Over het algemeen werden in vroegere trials grotere effecten geregistreerd dan in latere. De meest recente trials op hun beurt hadden doorgaans een betere methodologische kwaliteit. De beschikbare informatie suggereert nadrukkelijk dat Ginkgo effectief is bij patiënten met dementie en

andere vormen van cerebrale insufficiëntie. Slechts een beperkt aantal trials voldoet echter aan de huidige eisen ten aanzien van evaluatie en rapportage. Bovendien blijft het moeilijk de gerapporteerde effecten op hun juiste - klinische - waarde te schatten.

Hoofdstuk 3 schetst de opzet van de Maastrichtse Ginkgo Trial. De hoofdvraagstelling die het uitgangspunt vormde van deze interventiestudie luidde of het speciale, gestandaardiseerde *Ginkgo biloba* extract EGb 761 effectief is in de behandeling van oudere mensen met lichte tot matige dementie, respectievelijk met aan de leeftijd gerelateerd geheugenverlies (age-associated memory impairment, AAMI). Aanvullende onderzoeksvragen hadden betrekking op de relatie van enig effect met de dosis en de duur van de behandeling. Het onderzoek werd opgezet als een gerandomiseerde, placebo-gecontroleerde, dubbel-blinde, parallelle trial.

De deelnemers aan het onderzoek werden tussen januari 1994 en mei 1996 geworven in verzorgingshuizen in het zuiden van Nederland (Limburg, Zuid-Oost Brabant), op basis van een gefaseerde screeningsprocedure. Deelname aan het onderzoek vereiste in de eerste plaats de aanwezigheid van een licht of matig stadium van dementie (Alzheimer's dementie (AD), vasculaire dementie (VD), of een mengvorm van beide), of van leeftijdsafhankelijk geheugenverlies (AAMI). De aanwezigheid van een dementieel syndroom werd vastgesteld aan de hand van de SIDAM, een gestructureerd screeningsinterview, dat mede gebruik maakt van klinische beoordeling - in dit geval door de staf van het verzorgingshuis - en dat de DSM-III-R (thans: DSM-IV) en de ICD-10 criteria voor dementie op de voet volgt. In combinatie met de SIDAM werd de Syndrom Kurz Test (SKT) gebruikt om de mate van cognitieve achteruitgang vast te stellen. De SKT is een in Duitsland ontwikkelde maar ook in het En-

gelse taalgebied gevalideerde testbatterij met het accent op het meten van geheugen en aandacht (concentratie), bestaande uit 9 korte subtesten die elk maximaal 1 minuut in beslag nemen. Een score van 8-23 op de SKT (bereik: 0-27, met 27 als meest ongunstige waarde) gaf toegang tot de trial. Niet-demente ouderen werden als AAMI-patiënt geïdentificeerd op basis van een objectieve psychometrische test - andermaal de SKT, scores 8-23 - en een subjectieve beoordeling van de invloed van geheugenproblemen op het dagelijks leven (score ≥ 11 op de gemodificeerde Memory Assessment Clinics Questionnaire (MAC-Q), volgens de verzorging of de onderzoekspersoon zelf). Andere toelatingseisen hadden betrekking op de leeftijd (50 jaar of ouder), de bereidheid om schriftelijk in te stemmen met deelname aan het onderzoek, afwezigheid van een ernstige depressie (Geriatric Depression Scale (GDS), 15-item versie, score < 11), afwezigheid van ernstige co-morbiditeit en interfererende co-interventies, geen placebo-respons en voldoende compliance tijdens de inloopperiode van 3 weken die aan de start van de eigenlijke interventieperiode voorafging. Uiteindelijk werden 214 personen - 63 met dementie en 151 met AAMI - gerekruteerd in 39 verzorgingshuizen.

De behandeling bestond uit het slikken van tabletten die 0 mg (placebo), 80 mg, of 120 mg van het *Ginkgo biloba* extract EGb 761 bevatten. Speciale aandacht werd besteed aan de proefondervindelijke ontwikkeling van een hoogwaardig placebo-preparaat. Aan de deelnemers werden dagelijks twee tabletten verstrekt ('s morgens en 's avonds, in combinatie met de maaltijd). Toediening van de medicatie gebeurde onder toezicht van het verzorgend personeel. De verpakking en labelling van de onderzoeksmedicatie waren in overeenstemming met de geldende richtlijnen en wettelijke bepalingen. Toewijzing aan de alternatieve behandelingen (Ginkgo 240 mg per dag; Ginkgo 160 mg

per dag; placebo) gebeurde op basis van het lot, aan het begin van de interventieperiode en nog eens halverwege, na 12 maanden follow-up. De eerste randomisatie-ronde volgde onmiddellijk op de baseline-metingen. Hierbij werd een 2 : 2 : 1 toewijzingsratio gehanteerd. De randomisatie geschiedde binnen vooraf gedefinieerde strata op basis van diagnostische subcategorie (dementie vs. AAMI) en de gesignaleerde verandering in het geheugen tijdens de 6 maanden voorafgaand aan de interventie (verbetering of geen verandering vs. verslechtering). Uitsluitend deelnemers die de eerste 12 weken van de interventie voltooiden ($n=201$), kwamen in aanmerking voor de tweede randomisatie-ronde. De tweede randomisatie-ronde, waarbij het lot besliste over het al dan niet voortzetten van de Ginkgo-behandeling, was bedoeld om de persistentie van een Ginkgo-effect na het staken van de behandeling te kunnen evalueren.

Voorafgaand aan de start van de interventie werden diverse kenmerken gemeten die mogelijk de prognose van dementie en cognitieve stoornissen zouden kunnen beïnvloeden, zoals de leeftijd, het geslacht, en de aanwezigheid in de voorgeschiedenis van specifieke biologische 'life events' (BLEs: narcose, hersenschudding, etc.).

Om te toetsen of behandeling met Ginkgo het beoogde effect had, werden in het onderzoek verschillende meetinstrumenten gebruikt. Drie ervan werden vooraf als hoofdmaat aangewezen: de conclusies ten aanzien van de werkzaamheid van Ginkgo zouden in de eerste plaats op de resultaten met deze instrumenten gebaseerd worden. Deze primaire uitkomstmaten hadden betrekking op drie verschillende niveaus waarop een verminderd cognitief functioneren zich kan manifesteren: de prestaties op een neuropsychologische functietest (psychometrie), het globale klinisch beeld dat de ontwikkeling van de cognitieve stoornis oproept (globale beoordeling), en het waarneembare

gedrag dat met een betere of een slechtere cognitie samenhangt (gedragsobservatie). De gehanteerde beoordelingscriteria waren in overeenstemming met de meest recente richtlijnen voor de evaluatie van geneesmiddelen tegen dementie en cognitie-versterkende middelen. Naast de primaire uitkomstmaten werden verschillende andere uitkomstmaten gebruikt, bijvoorbeeld maten die in eerder onderzoek naar de effectiviteit van *Ginkgo biloba* gebruikt waren en tot een positieve uitkomst hadden geleid, en maten die wellicht op een specifieke manier iets zouden kunnen laten zien van de manier waarop Ginkgo de hersenfunctie beïnvloedt. Deze maten waren echter van secundair belang bij het trekken van conclusies omtrent de werkzaamheid van Ginkgo.

De volgende instrumenten fungeerden als primaire uitkomstmaat. 1. De Syndrom Kurz Test (SKT), een korte testbatterij met het accent op geheugen en aandacht (score-bereik: 0-27), die tevens werd gebruikt als screeningstest om kandidaten voor deelname aan het onderzoek te identificeren. De SKT werd in het kader van de follow-up interviews afgenomen door getrainde interviewers. 2. Item 2 van de 'Clinical Global Impression' (CGI-2), dat berust op een globale hetero-beoordeling van de waargenomen verandering in de cognitief functioneren sinds het begin van de interventie. In de Maastrichtse Ginkgo Trial werd de CGI (7 punt schaal; 1 = zeer sterke verbetering; 4 = geen verandering; 7 = zeer sterke verslechtering) gescoord door het meest direct betrokken lid van het verzorgend personeel van de instelling waarin de onderzoekspersoon verbleef. 3. De Nürnberger Alltags-Aktivitäten Skala (NAA), een eveneens in Duitsland ontwikkelde beoordelingslijst van een 20-tal cognitie-afhankelijke dagelijkse levensverrichtingen op een instrumenteel niveau (score-bereik: 20-60). Deze schaal vormt onderdeel van de Nürnberger Alters Inventar (NAI), een verzameling vragenlijsts-

ten, beoordelingsschalen en psychometrische tests die specifiek gericht is op toepassing bij oudere mensen. Tot de secundaire uitkomstmaten behoorden in de eerste plaats diverse psychometrische testen: een getallen-verbindings-test (NAI-ZVT-G, 'trail making test'), een test op de mondelinge reproductie van cijferreeksen (NAI-ZN-G, 'digit span test'), en een test gericht op het meten van de onmiddellijke reproductie en de vertraagde herkenning van woorden (NAI-WL, 'word list'). Daarnaast werd gebruik gemaakt van enkele klinische beoordelingsschalen: een lijst met 18 geriatrische symptomen (SCAG = Sandoz Clinical Assessment-Geriatric scale), een lijst met depressie-verschijnselen (GDS = Geriatric Depression Scale, verkorte 15-item versie), de zelf-gepercipieerde gezondheid (rapportcijfer), en zelf-gepercipieerde geheugenfunctie (rapportcijfer). Bovendien werd de NAI-NAA ook door de deelnemers zelf gescoord. Voor het begin van de interventie werden van alle uitkomstmaten baseline-gegevens verzameld. Alle uitkomstparameters werden gemeten na een behandelduur van 12 en 24 weken. Een deel van de uitkomstmaten (SKT, CGI-2, SCAG, gepercipieerde gezondheid, gepercipieerde geheugenfunctie) werd bovendien tussentijds geëvalueerd, na 4, 8 en 18 weken follow-up. Behalve de beoogde uitkomsten werden op ieder meetpunt (4, 8, 12, 18, en 24 weken) diverse andere relevante verschijnselen gemeten, namelijk het optreden van emotionele gebeurtenissen die mogelijk invloed zouden kunnen hebben op het cognitief functioneren (PLEs = Psychological Life Events), de mate van therapietrouw (tellen van resterende tabletten), en het optreden van (ernstige) ongewenste gebeurtenissen. Dit laatste gebeurde aan de hand van een checklist van gangbare gezondheidsklachten en op basis van spontane rapportage van eventuele complicaties door het verzorgend personeel, de bewoners zelf, of hun familie. Om uit te kunnen

sluiten dat het stoppen met Ginkgo onthoudingsverschijnselen oproept, werd na 18 en 24 weken follow-up een checklist met een aantal bij diverse vormen van verslaving gangbare klachten beoordeeld. Na 12 en 24 weken werd voorts de blootstelling aan bepaalde co-interventies geregistreerd (co-medicatie, verandering in de zorg). Om inzicht te krijgen in de beoogde en de niet-voorzienbare behandel-effecten werden in essentie twee verschillende methoden van gegevensverzameling gebruikt: interviews met de onderzoeksdeelnemers door een team van getrainde interviewers; en observaties door het verzorgend personeel in elk van de participerende huizen.

De hoofdanalyse van de onderzoeksuitkomsten bestond uit een 'intention-to-treat' analyse van de (veranderingen in) de scores op de primaire uitkomstmaten na 24 weken bij die deelnemers die gedurende de hele interventieperiode Ginkgo of placebo hadden gekregen. Testen voor de afzonderlijke uitkomstmaten werden voorafgegaan door een overall OLS-test van het effect op uitkomstmaten samen. Ook werd een responderanalyse uitgevoerd, die zich richtte op een vergelijking van het aantal onderzoekspersonen met een klinisch relevant effect op minstens twee van de drie primaire uitkomstmaten. In aanvulling op deze confirmatorische analyses, bedoeld om conclusies te trekken over de effectiviteit van Ginkgo, werd een groot aantal descriptieve analyses verricht, om inzicht te krijgen in de reacties op de secundaire uitkomstmaten, de effecten na een behandelduur van 12 weken, de dosis-effect relatie, het verschil in effect tussen een langere (24 weken) en een kortere (12 weken) behandelduur, de effecten bij subgroepen met specifieke kenmerken, en de effecten bij uitsluitend die onderzoekspersonen die niet noemenswaardig van het behandelprotocol waren afgeweken ('per-protocol' analyse, 'valid cases' analyse).

In **hoofdstuk 4** worden de belangrijkste uitkomsten van de Maastrichtse Ginkgo Trial beschreven. Dit hoofdstuk beperkt zich tot de resultaten op basis van de primaire uitkomstmaten (SKT, CGI-2, NAI-NAA). Via de trapsgewijze screeningsprocedure werden uiteindelijk in de 39 deelnemende verzorgingshuizen 214 proefpersonen gevonden die bereid waren aan het onderzoek mee te doen, en die bovendien aan alle andere toelatingseisen voldeden. De gemiddelde leeftijd van de gerandomiseerde populatie was 82,6 jaar; de populatie bestond uit 180 vrouwen en 34 mannen; 63 bewoners werden als dement geassocieerd, terwijl bij de overige 151 weliswaar sprake was van een opmerkelijk geheugen- en cognitieverlies (AAMI = Age-Associated Memory Impairment, oftewel leeftijd-gebonden achteruitgang van het geheugen), echter zonder dat aan alle overige criteria voor dementie voldaan was.

Van de 214 deelnemers aan het onderzoek vielen er 18 uit voor het einde van de rit (24 weken), 13 gedurende de eerste helft van de interventie (weken 0-12), en nog eens 5 gedurende de tweede helft van de interventie (weken 13-24). In de meeste gevallen was het overlijden of een ernstige ziekte van de onderzoekspersoon de aanleiding voor het voortijdig afbreken van de behandeling. In de groepen die aanvankelijk met Ginkgo 240 mg, Ginkgo 160 mg, en placebo waren behandeld, werden respectievelijk 11%, 6%, en 8% uitvallers geregistreerd. Deze verschillen hebben niet veel om het lijf. Gedurende de totale looptijd van de trial bleken voorts 28 personen - onder wie de zojuist genoemde uitvallers - zich onvoldoende te hebben gehouden aan het behandelprotocol (inname van minimaal 75% van de voorgeschreven onderzoeksmedicatie, metingen ongeveer op de geplande data). Ook deze 'non-compliers' waren vrijwel gelijk verdeeld over de onderscheiden behandelgroepen. Bij aanvang van de interventie werden tussen de behandelgroepen voor geen van de

onderzochte kenmerken statistisch significante of anderszins opvallende verschillen gevonden. Vier weken na de start van de behandeling werd aan alle betrokken beoordelaars - personeel van de verzorgingshuizen, interviewers, onderzoekspersonen zelf - gevraagd een uitspraak te doen over de vermeende aard van de ingestelde behandeling (Ginkgo 240, Ginkgo 160, of placebo). De analyse van de betreffende gegevens wees uit dat van een systematische overeenstemming tussen de waargenomen en de feitelijke behandeling geen sprake was. Een herhaling van deze 'check op blinding', in week 18 - 6 weken na de start van de tweede interventie-periode - leverde een soortgelijke uitslag op. Op grond van deze resultaten vertrouwen we erop dat de uitkomstmetingen in het kader van de studie inderdaad onder 'blinde' condities zijn geschied, en dat de meetresultaten niet beïnvloed zijn door voor-kennis omtrent de aard van de onderzoeksmedicatie.

Uiteindelijk werden 123 van de 214 gerandomiseerde onderzoekspersonen toegelaten tot de hoofdanalyse van de onderzoeksresultaten ('confirmatorische analyse'). Het ging hierbij om die personen die gedurende 24 weken uitsluitend Ginkgo ($n=79$, van wie 39 in een dosering van 240 mg per dag en 40 in een dosering van 160 mg per dag), of uitsluitend placebo ($n = 44$) hadden gebruikt, en voor wie bovendien meetresultaten bij het begin van de interventie en na 24 weken follow-up beschikbaar waren. 78 Personen die gedurende het eerste deel van de trial aan een Ginkgo-behandeling (38 in een dosering van 240 mg per dag en 40 in een dosering van 160 mg per dag) en gedurende het tweede deel aan een placebo-behandeling waren toegewezen, kwamen niet in aanmerking voor deze analyse. Een statistische overall toets waarbij de drie primaire uitkomstmaten (SKT, CGI-2, NAI-NAA) gelijktijdig in beschouwing genomen werden, liet geen effect

van Ginkgo zien ($T_{OLS} = 0.34$; $p = 0.37$). Ook voor de primaire uitkomstmaten apart werd geen effect gevonden. Geheugen en aandacht vertoonden ten opzichte van de start van het onderzoek een geringe achteruitgang in zowel de Ginkgo-groep (240 mg/dag en 160 mg/dag gecombineerd) als de placebogroep: $-0,82$ vs. $-1,23$ punten op de SKT-schaal, die loopt van -27 (maximaal haalbare verslechtering) tot $+27$ (maximaal haalbare verbetering). Gecorrigeerd voor de geringe verschillen in kenmerken bij de baseline-meting bedroeg het effect $+0,41$ (90% betrouwbaarheidsinterval: $-0,86 - +1,68$), ten gunste van de Ginkgo-groep. Ook de cognitieve afhankelijke dagelijkse levensverrichtingen bleken in beide groepen achteruit te zijn gegaan: $-1,37$ vs $-1,36$ punten op de NAI-NAA schaal, die een bereik heeft van -40 (maximale verslechtering) tot $+40$ (maximale verbetering). Het voor confounding gecorrigeerde verschil bedroeg $-0,36$ (90% betrouwbaarheidsinterval: $-1,94 - +1,22$), waarbij de negatieve waarde duidt op een beter resultaat in de placebogroep. De globale beoordeling van de cognitieve verandering door de zorgfunctionarissen in het verzorgingstehuis leverde evenmin een relevant effect op: $4,29$ vs $4,34$ op de CGI-schaal, die loopt van 1 (zeer sterke verbetering), via 4 (geen verandering), tot 7 (zeer sterke verslechtering), oftewel een effect van $+0,05$ (95% betrouwbaarheidsinterval: $-0,31 - +0,41$) ten faveure van de Ginkgo-groep.

Bij de analyse van de interventie-effecten na 12 weken konden meer onderzoekspersonen betrokken worden: 166 in de beide Ginkgo groepen samen (82 met een dagelijkse dosis van 240 mg, en 84 met een dagelijkse dosis van 160 mg), en 48 in de placebo-groep. Ook deze analyse liet echter geen noemenswaardig effect van Ginkgo zien, noch op de SKT ($+1,00$; 90% betrouwbaarheidsinterval: $-0,05 - +2,05$), noch op de NAI-NAA ($+0,25$; 90% betrouwbaarheidsin-

terval: $-0,98 - +1,48$), noch op de CGI-2 ($-0,03$; 95% betrouwbaarheidsinterval: $-0,27 - +0,21$). Na 24 weken vertoonde 14% van de personen in de beide Ginkgo-groepen samen en 18% van de personen in de placebo-groep een klinisch relevante respons op de behandeling, dat wil zeggen dat zij voldeden aan minimaal twee van de volgende criteria: verbetering op de SKT-schaal met minimaal 3 punten; verbetering op de NAI-NAA schaal met minimaal 4 punten; en een score <4 op de CGI-2. Dit verschil was statistisch niet significant ($\chi^2 = 0,39$; $p = 0,53$). Noch de patiënten met dementie, noch de AAMI-patiënten, noch de patiënten behorende tot enige andere subgroep binnen de onderzoekspopulatie hadden merkbaar baat bij Ginkgo. 240 Mg Ginkgo per dag werkte niet beter dan 160 mg per dag. Evenmin leidde doorgaan met Ginkgo tot een beter resultaat dan stoppen na 12 weken. Ongewenste gebeurtenissen, al dan niet van ernstige aard, traden in de Ginkgo-groep niet significant vaker op dan in de placebo-groep. Een kritische reflectie op de opzet en uitvoering van de Maastrichtse Ginkgo Trial, in het bijzonder met betrekking tot de juistheid van de keuze en de vergelijkbaarheid van de onderzoekspopulatie, de interventiecondities en de metingen, leverde geen afdoende verklaring op voor de afwezigheid van enig effect.

In **hoofdstuk 5** worden de resultaten van de Maastrichtse Ginkgo Trial voor een aantal aanvullende uitkomstmaten beschreven. Evenals de primaire uitkomstmaten waren deze secundaire uitkomstmaten gebaseerd op psychometrisch onderzoek, klinische beoordeling, en gedragsbeoordeling. Na onafgebroken behandeling met Ginkgo ($n=79$), dan wel placebo ($n=44$) gedurende 24 weken viel voor geen enkele van deze maten een statistisch significant en betekenisvol verschil in gemiddelde score-verandering tussen de Ginkgo-groep (240 en 160 mg/dag

gecombineerd) en de placebogroep te noteren. Na correctie voor baseline-verschillen vonden we de volgende effecten, waarbij een positieve waarde in het voordeel van Ginkgo spreekt: getallen-verbindingstest (NAI-ZVT-G): $+18,4$ (90% betrouwbaarheidsinterval: $-0,7 - +37,5$) op een schaal van -300 (maximale verslechtering) tot $+300$ (maximale verbetering); mondelinge reproductie van cijferreeksen (NAI-ZN-G): $-0,01$ (90% betrouwbaarheidsinterval: $-0,54 - +0,52$) op een schaal van -17 (maximale verslechtering) tot $+17$ (maximale verbetering); reproductie en herkenning van een woordenlijst (NAI-WL): $+0,12$ (90% betrouwbaarheidsinterval: $-0,69 - +0,93$) op een schaal van -16 (maximale verslechtering) tot $+16$ (maximale verbetering); geriatrische symptomen (SCAG): $+0,01$ (90% betrouwbaarheidsinterval: $-0,22 - +0,24$) op een schaal van $-6,0$ (maximale verslechtering) tot $+6,0$ (maximale verbetering); depressieve stemming: $+0,54$ (90% betrouwbaarheidsinterval: $-0,64 - +1,54$) op een schaal van $-15,0$ (maximale verslechtering) tot $+15,0$ (maximale verbetering); gepercipieerde gezondheidstoestand (rapportcijfer): $-0,15$ (90% betrouwbaarheidsinterval: $-0,56 - +0,26$) op een schaal van $-10,0$ (maximale verslechtering) tot $+10,0$ (maximale verbetering); gepercipieerde geheugenfunctie (rapportcijfer): $-0,15$ (90% betrouwbaarheidsinterval: $-0,51 - +0,21$) op een schaal van $-10,0$ (maximale verslechtering) tot $+10,0$ (maximale verbetering); en zelf-gepercipieerd niveau van cognitieve-dagelijkse levensverrichtingen (NAI-NAA, gescoord door onderzoekspersoon zelf): $+1,32$ (90% betrouwbaarheidsinterval: $+0,13 - +2,77$) op een schaal van -40 (maximale verslechtering) tot $+40,0$ (maximale verbetering). Na 12 weken deden de personen in de Ginkgo-groep ($n=166$; beide doseringen samengevoegd) het een beetje beter dan de personen in de placebo-groep met betrekking tot de instrumentele dagelijk-

se levensverrichtingen (NAI-NAA: +1,35 (90% betrouwbaarheidsinterval: +0,22 - +2,49)), maar een beetje slechter ten aanzien van de zelf-gepercipieerde gezondheid (-0,38 (90% betrouwbaarheidsinterval: -0,73 - -0,04)). Een relatief hoge Ginkgo dosis (240 mg per dag) werkte niet beter dan de gangbare dosis (160 mg per dag). Ook prolongatie van de behandeling na 12 weken zette geen zoden aan de dijk. We slaagden er niet in enige subgroep te identificeren die systematisch baat had bij Ginkgo. De resultaten van een 'per-protocol' analyse, waarin alleen die personen betrokken werden die in alle opzichten voldaan hadden aan de eisen van het interventie-protocol, gaf in vergelijking met de initiële 'intention-to-treat' analyse geen afwijkende resultaten te zien. De resultaten van de analyse op basis van de secundaire uitkomstmaten bleken congruent met die van de primaire uitkomstmaten (hoofdstuk 4). De conclusie naar aanleiding van de resultaten van de Maastrichtse Ginkgo Trial was dan ook dat het speciale *Ginkgo biloba* extract EGb 761 niet effectief was gebleken bij ouderen met een lichte tot matige vorm van dementie, of met een substantiële reductie van de geheugenfunctie.

Hoofdstuk 6 heeft het karakter van een algemene discussie, waarbij getracht wordt de opzet en de resultaten van de Maastrichtse Ginkgo Trial te confronteren met die van eerdere Ginkgo-trials in het domein van cerebrale insufficiëntie, cognitieve achteruitgang en dementie. Op grond van een kritische beschouwing van de trials die gepubliceerd zijn voorafgaand aan dan wel na de start van het Maastrichtse experiment moet geconcludeerd worden dat een substantiële bewijslast ten grondslag ligt aan de claim dat Ginkgo heilzaam is bij een (dreigende) vermindering van de hersenfunctie. Niettemin lijkt deze bewijslast voor het grootste deel zijn oorsprong te vinden in een klein aantal trials van recente datum, die zich richtten op pa-

tiënten met dementie, en die lijken te zijn opgezet conform de meest actuele richtlijnen voor de evaluatie van noötropica en geneesmiddelen tegen dementie. In vergelijking met de grootte van effecten gerapporteerd uit eerdere studies, laten de meer recente trials naar verhouding vrij bescheiden effecten zien. Vooralnog zijn er geen concrete aanwijzingen dat de consistente uitkomsten van eerdere trials wat betreft de richting van het effect van *Ginkgo biloba*, het gevolg zouden kunnen zijn van een selectief publicatie-beleid. Daar staat echter tegenover dat de meeste onderzoeksactiviteiten op dit terrein in het verleden zijn uitgevoerd zonder dat er sprake was van strikte en nauwgezette supervisie door een onafhankelijke instantie. Ten aanzien van de Maastrichtse Ginkgo Trial zijn we de mening toegedaan dat met name de randomisatie-procedure, de gerealiseerde mate van therapie-trouw, het scherpe interventie-contrast op basis van een hoogwaardig placebo-preparaat, en het mede daarop geënte 'blinde' karakter van de uitkomstmetingen hebben bijgedragen aan de wetenschappelijke rigueur van de trial. Meer twijfel is op zijn plaats ten aanzien van de keuze van de uitkomstmaten en de uitvoering van de metingen, in combinatie met de keuze van de onderzoekssetting en de samenstelling van de onderzoekspopulatie. Vooral op pragmatische gronden - beschikbaarheid van onderzoekskandidaten en relevante gegevens, gemakkelijke organisatie, bevorderen van compliance, mogelijkheid tot hetero-beoordeling door het verzorgend personeel - is gekozen voor verzorgingshuizen als bron voor rekrutering van de patiëntenpopulatie. Keerzijden van deze wervingsstrategie waren de hoge gemiddelde leeftijd van de onderzoekspopulatie, de hoge mate van co-morbiditeit, en een situatie waarin een verval van cognitieve functies relatief gemakkelijk gecompenseerd wordt. Voorzover de keuze en het gebruik van meetinstrumenten en deelnemers al tot ver-

tekening van de onderzoeksresultaten aanleiding heeft gegeven, is het aannemelijk dat het hierbij gaat om non-differentiële mechanismen, resulterend in een (gedeeltelijke) onderdrukking van een eventueel aanwezig interventie-effect. Mede gelet op de grote variatie in gemeten kenmerken, gebruikte meetinstrumenten en ingezette waarneemers, zou het echter vreemd zijn indien een reëel bestaand effect door de gesuggereerde tekortkomingen geheel aan het oog onttrokken zou worden.

Ondanks aantoonbare verschillen in de onderzoeksopzet, in het bijzonder ten aanzien van de samenstelling van de onderzoekspopulatie, is het ons onvoldoende gelukt de resultaten van de Maastrichtse Ginkgo Trial met die van eerdere studies te verzoenen. We ontkomen dan ook niet aan de conclusie dat één of meer nieuwe, grootschalige, zorgvuldig opgezette trials nodig zijn om meer duidelijkheid te krijgen over de effectiviteit van *Ginkgo biloba* bij dementie en cognitieverlies, ook al klinkt deze conclusie in het licht van de oorspronkelijke doelstelling van de Maastrichtse trial wellicht wat goedkoop en obligaats.

DANKWOORD

Toen ik de gelegenheid kreeg met het onderzoek naar Ginkgo te beginnen, dacht ik: bingo!! Overigens zonder te beseffen dat dit gezelschapsspel - in Limburg ook wel kienen genoemd, en wellicht niet geheel ten onrechte - tijdens onze studie in meer dan één opzicht een serieuze concurrent dreigde te worden van de hoofdrolspeler: Ginkgo. De berichten over *Ginkgo biloba* waren veelbelovend, spectaculair bijna. Van Ginkgo zou je kien worden! Tenminste, als je uitkomsten van eerder onderzoek mocht geloven. Nou, daar wil je natuurlijk graag het fijne van weten. Reeds tijdens de voorbereidingen van de trial rezen bij mij enige twijfels. Naarmate het onderzoek vorderde, werden die alleen maar sterker. Ik was me ervan bewust dat we met iets knap lastigs bezig waren. "Meten is Weten" is een wijsheid waar ik nooit erg van gecharmeerd ben geweest. Uiteindelijk bleek de uitkomst van de Maastrichtse Ginkgo Trial negatief te zijn: Ginkgo 'deed' niets. Misschien had deze uitslag mij als onafhankelijk onderzoeker onberoerd moeten laten. De waarheid aan het licht brengen, dat is immers waar het in de wetenschap om draait. Toch heb ik het negatieve resultaat van ons onderzoek als een behoorlijke domper ervaren, om verschillende redenen. Vooral het onvermogen een verdedigbare verklaring voor onze afwijkende onderzoeksresultaten te vinden werkte verlamdend en heeft een vlotte afhandeling van het onderzoek - met inbegrip van de academische verdediging - in de weg gestaan. Niet dat ik dit Ginkgo-onderzoek als mijn levenswerk beschouw, zeker niet. Maar toch begon ik mij stilaan af te

vragen: promoveer ik nog wel vóór ik grijs (of erger) ben? Welnu, het zal erom spannen. 26 Maart zal het leren. Er kan soms in een week veel gebeuren. Soms ook niet.

We hebben buitengewoon veel plezier beleefd aan het veldwerk en de contacten met onze proefpersonen. Een eerste blijk van dank is dan ook op zijn plaats voor de bewoners van de verzorgingshuizen en hun familie. De bereidheid tot medewerking en de getoonde belangstelling voor het onderzoek waren fenomenaal. Ik had er graag wat meer voor teruggegeven.

Ook de directies, de hoofden zorg, de afdelingshoofden, en de rest van het verzorgend personeel in de deelnemende verzorgingshuizen zijn we bijzonder erkentelijk.¹ De coöperativiteit en getoonde inzet waren boven verwachting. Aan jullie heeft het beslist niet gelegen dat we er niet in zijn geslaagd het beoogde aantal van 300 deelnemers te recruter en dat de teller bij 214 is blijven steken. Ik realiseer me het in ons gestelde vertrouwen enigszins beschaamd te hebben door in een zo laat stadium over de onderzoeksuitkomsten te rapporteren. De belangrijkste oorzaak is bekend: we bleken niet bij machte te zijn een ondubbelzinnig antwoord te geven op de hamvraag ('Werkt *Ginkgo biloba*?'). Verder hoop ik dat met het verschijnen van dit proefschrift niet de onjuiste indruk wordt gewekt dat de daaraan gekoppelde academische promotie het hoofdmotief was om de Maastrichtse Ginkgo Trial te initiëren. Een verbetering van de zorg voor ouderen, dat stond ons primair voor ogen.

Jos Kleijnen, jij stond aan de wieg van de Maastrichtse Ginkgo Trial, vooral met je eerdere review-artikelen over Ginkgo, onder andere in de Lancet. De subsidie voor het onderzoek was reeds geregeld voordat een in detail uitgewerkte onderzoeksopzet en begroting opgesteld waren. Dankzij jou kwam ik in een gespreid bedje terecht. In de loop der jaren ben ik steeds jaloerser geworden op de soepele en doortastende manier waarop jij literatuuroverzichten genereert en dissemineert. Nog vóór de eigenlijke start van het onderzoek verliet je ons team om je carrière in Amsterdam een vervolg te geven.

Paul Knipschild, nu mijn promotor, destijds was jij samen met Jos de initiator van de Maastrichtse Ginkgo Trial. Het design van het onderzoek draagt nadrukkelijk jouw stempel. In de loop van het project ben je er noodgedwongen een tijdje tussenuit geweest, om ruim vóór de afronding van het project in volle glorie terug te keren, onder meer met adviezen op weer een ander terrein dat jij als geen ander beheerst: de begrijpelijke verslaglegging van onderzoek. Omdat ik misschien wel net zo eigenwijs ben als jij, heb ik je heel wat van mijn water in jouw wijn laten doen en is het uiteindelijk een ander proefschrift geworden dan jou wellicht voor ogen stond. Minder scherp, minder spits, minder sprankelend ook. Ik heb in de loop van de jaren enorm veel waardering gekregen voor de kritische en intensieve wijze waarop jij onderzoek en onderzoekers begeleidt. Toen ik je laatst vroeg of ik je nog niet op je zenuwen begon te werken, antwoordde je ontkennend. Maar misschien was dat bluff. Had de zweep er bij mij wellicht steviger overheen gemoeten? Ik betwijfel of dat het gewenste resultaat zou hebben gehad!

Hilde Sielhorst en Erik van Rossum, samen met mij vormden jullie de harde kern van het Ginkgo-team.

Erik, na het vertrek van Jos nam jij zijn plaats in. Wij vulden en voelden elkaar goed aan, althans in mijn perceptie. Ik maakte de knopen, jij hakte ze door. Voorbeeldig teamwork. Dat het Ginkgo-onderzoek de eindstreep heeft gehaald, is voor een belangrijk deel jouw verdienste. Je deed echter meer. Je stelde alles in het werk om mij zo snel mogelijk weer terug te brengen op het rechte pad, zodra ik weer eens vreemd dreigde te gaan. Je fungeerde regelmatig als buffer, zo niet als bodyguard, tussen mij en de soms ietwat bozige buitenwereld. Jouw geheim: je beschikt, denk ik, over een uniek talent om met mensen om te gaan. Maar wee diegene die deze ontboezeming opvatten als een poging jouw wetenschappelijke capaciteiten en organisatievermogen te bagatelliseren. Die krijgt het met mij aan de stok!

Hilde, jij vormde het kloppende hart van het onderzoek, het bindmiddel, het zenuwcentrum waar alle in- en uitgaande signalen samen kwamen. Als jij op het toneel verscheen, met of zonder vlaaien, dan gingen alle deuren open. Op vele aspecten van het onderzoek had jij van ons allemaal verreweg de meeste kijk, daar ben ik van overtuigd. Het is nu eenmaal mijn beurt, anders had ik op het podium graag met je van plaats geruild. Opmerkelijk is verder dat je al die jaren nooit merkbaar boos op mij geworden bent, hoewel ik daartoe regelmatig aanleiding moet hebben gegeven.

Fons Kessels, net als bij vele andere en eerdere projecten was jij ook bij de Ginkgo trial weer het manusje-van-alles, het schaaftmet-de-vijf-poten, het duizend-dingen-doekje. Maar zeker niet het eerste het beste! Je randomiseerde, je etiketteerde, je analyseerde, je chauffeerde ons naar Karlsruhe, en je serveerde lollies dat het een lieve lust was. Je fungeerde bovendien als trial-monitor, en je had de hele Ginkgo-voorraad bij je thuis te logeren. Wie eraan twijfelt of dat al-

les wel samen door één deur kan: Fons vormt het levende bewijs. Als wij het niet meer snapten, konden we altijd bij jou aankloppen. Je adviezen en ingevingen waren soms zo geniaal dat ik je ervan verdacht thuis stiekem van de voorraad te snoepen. Totdat je - uiteraard weer jij - de randomisatie-code verbrak

Dr. Robert Hörr, Dr. Günther Meng en Dr. Meinhard Kieser, ik ben zeer onder de indruk geraakt door de professionele wijze waarop jullie de Maastrichtse Ginkgo Trial namens de sponsor, Dr. Willmar Schwabe Arzneimittel GmbH & Co in Karlsruhe, hebben begeleid. Ondanks de grote belangen die voor jullie op het spel stonden is de samenwerking van het begin tot het eind correct, ontspannen en plezierig verlopen. Wij hebben er enorm veel van geleerd. Alleen al om die reden vond ik het een moeilijke opdracht jullie van de uitkomsten van onze trial op de hoogte te stellen.

Marijke Moll, jij assisteerde het Ginkgo-team gedurende een substantieel deel van het project, met name bij de screeningsinterviews en diverse andere klussen in het kader van de gegevensverzameling en dataverwerking. De manier waarop je dat deed - zeer betrokken, nauwgezet en degelijk - heeft diepe indruk op me gemaakt.

Sacha van de Crommert, jouw faam als data-manager en -koppelaarster, was je reeds vooruit gesnelde, toen je met de Ginkgo-bestanden in de weer ging. Desondanks heb je me toch nog verstedeld doen staan. Toen alle onderzoeksgegevens binnen waren en er geanalyseerd kon gaan worden, raakte ik binnen de kortste keren compleet confuus van die vele honderden variabelen, die ik nota bene zelf bedacht had. Als relatieve buitenstaander had jij aan een paar woorden genoeg om die weerbarstige massa in een mum van tijd in het gareel te krijgen.

Astrid Versantvoort en Elian Douven, als student-assistenten hebben jullie bergen werk verzet, vooral in perioden waarin de interview-verplichtingen ons boven het hoofd dreigden te groeien. Maar ook op andere terreinen hebben jullie je verdienstelijk gemaakt, bijvoorbeeld het ordenen en ontsluiten van de literatuur. Ik besef dat ik met jullie voorzet best wat meer had mogen doen.

Medewerksters van de Interview Service van het MEMIC, jullie hebben ons een geweldige dienst bewezen met de wijze waarop jullie de follow-up interviews met de deelnemers aan het onderzoek hebben uitgevoerd. De opdracht waarmee wij jullie in zee gestuurd hebben, was allesbehalve eenvoudig. De enthousiaste verhalen waarmee jullie ons tussentijds van jullie ervaringen in kennis stelden, maakten duidelijk dat jullie er alles aan hebben gedaan om het onderzoek tot een succes te maken.

Cobie Martens, alweer een hele tijd geleden zat jij gretig in de startblokken om van mijn proefschrift iets fraais te gaan maken. Zolang ik nog zorgen had over de inhoud, kon ik me echter slecht op de vormgeving concentreren. En toen dat laatste toch moest, toonde ik me ook in dat opzicht nogal wankelmoedig. Ik heb je daardoor de kans ontnomen je talenten op dit gebied volop te etaleren. Dat spijt me oprecht. Op het laatst waren we allebei behoorlijk zenuwachtig. Van mijn kant is dat begrijpelijk, jou siert het. Het had voorwaar niet veel gescheeld, of je had *mij* in twee kolommen ingedeeld! Gelukkig hebben we op de valreep in overleg met Datawysse een aanvaardbare oplossing weten te vinden.

Pieter Dagnelie en Riekje de Vet, samen met Erik namen jullie zitting in de speciale 'task force' ter verhoging van de Ginkgo-productiviteit. Ik kan me voorstellen dat jullie zelf met het resultaat van jullie interventie niet onver-

deeld gelukkig zijn. Maar, geloof me, het heeft beslist geholpen.

Piet van den Brandt, als vakgroepvoorzitter en aandachtsveldleider voedingsepidemiologie heb ik het je de afgelopen jaren niet gemakkelijk gemaakt. Het siert je dat je, misschien wel eens binnensmonds knarsend, altijd een constructieve houding bent blijven aannemen.

Ferd Sturmans, jij hield mij bij de hand toen ik mijn eerste schuchtere schreden in epidemiologieland zette, samen met Jos Baecke en Gerhard Zielhuis. Ik heb een tijdlang geprobeerd je sporen te volgen. Jouw tempo werd mij evenwel te machtig en ik heb moeten lossen.

Een speciaal woord van dank wil ik richten tot alle familieleden, vrienden en kennissen in 'het veld', onder wie mijn ouders, Floor en Rob, Lex, en diverse anderen. Door bij iedere gelegenheid naar de stand van zaken en liever nog de voortgang van mijn onderzoek te informeren, hebben zij blijk gegeven over een ijzersterk geheugen en een grote vasthoudendheid te beschikken. Ik hoop hen binnenkort weer onder ogen te kunnen komen. Yvonne, Ria, andere (ex-)medewerkers van het secretariaat, de jonkies binnen onze vak-

groep niet te vergeten. Die mysterieuze, vaak wat gedesoriënteerde man die jullie de afgelopen jaren soms aantroffen op de vakgroep, soms ook niet: dat was ik!

Beste mensen van de capaciteitsgroep Epidemiologie, hier in Maastricht. Het spijt mij dat jullie er op 26 maart helemaal alleen voor staan. Via enkele halsbrekende toeren op de valreep en door mij terug te trekken voor de retraite, heb ik geprobeerd jullie nog een uitdaging helpend handje toe te steken. Meer kon ik niet doen. Sterkte toegewenst met het offensief! Als ik dit overleef, ben ik de volgende keer beslist weer van de partij.

Wouter, Désirée, Sophie en Maria, ook jullie hebben de afgelopen tijd behoorlijk afgezien. Zelfs ons Ginkgo-boompje heeft deze barre periode niet overleefd. Ik zal daarom proberen mijn zolderkamertje de komende tijd wat vaker te laten voor wat het is.

Tot slot wil ik alle andere, niet met name genoemde personen bedanken die op enigerlei wijze een bijdrage hebben geleverd aan de totstandkoming van dit proefschrift, in het bijzonder degenen die er geen prijs op stellen daar op deze plaats open en bloot aan herinnerd te worden.

1. De volgende verzorgingshuizen participeerden in de Maastrichtse Ginkgo Trial:

Lenculenhof (Maastricht), Franciscus (Beek), Molenhof (Maastricht), Ave Maria (Geulle), KBO De Beyart (Maastricht), Providentia (Maastricht), Kruisberg (Brunssum), Olympiastaete (Geleen), Lindenheuvel (Geleen), Bunderhof (Geleen), Moutheuvel (Stein), De Schutse (Hoensbroek), Pius Centrum (Hoensbroek), Beukeloord (Meerssen), Damiaanberg (Meerssen), Ter Eyck (Heerlen), Douvenrade (Heerlen), De Regenboog (Heerlen), Op den Toren (Nuth), Ackenshuis (Gulpen), Aelserhof (Elsloo), 't Brook (Voerendaal), Tobias (Heerlen), St. Anna (Heerlen), Bocholtz (Bocholtz), St. Jan (Maasmechelen, B), Roncalli (Roermond), RCG (Roermond), Petrusberg (St Odiliënberg), Rust in Roy (Stramproy), De Erkenkamp (Steyl), Oppe Ruiver (Reuver), Bartholomeus (Asten), Bethanië (Mierlo), De Ruyschenberg (Gemert), Nieuwenhof (Mierlo), De Wolfskoele (Vaals), De Eeuwsels (Helmond), Elvira (Amstenrade).

CURICULUM VITAE

Martien van Dongen werd geboren op 24 december 1953 in Rijen, Noord-Brabant. In juni 1972 behaalde hij het Gymnasium-diploma aan het Mill Hill College in Tilburg. In datzelfde jaar begon hij met de studie Humane Voeding aan de Landbouw Hogeschool in Wageningen. Deze opleiding sloot hij in september 1980 cum laude af, met als doctoraalvakken Voeding en Gezondheidsleer/Epidemiologie. Een onderzoeks- en praktijkstage in Sri Lanka, waarbij de voedingstoestand van de theeplantage-arbeiders centraal stond, vormde een klein, maar door zijn vormend karakter erg belangrijk deel van deze opleiding. Van juni 1980 tot november 1981 was Martien werkzaam bij het Instituut van Sociale Geneeskunde van de Katholieke Universiteit Nijmegen, waar hij zich bezig hield met oriënterend onderzoek naar de mogelijkheden tot epidemiologisch onderzoek in de arbeids- en bedrijfsgezondheidszorg, alsmede met diverse onderwijstaken.

In november 1981 werd Martien als universitair docent aangesteld bij de capaciteitsgroep Epidemiologie/Gezondheidszorgonderzoek (thans: Epidemiologie) binnen de

Algemene Faculteit (thans: Faculteit der Gezondheidswetenschappen) van de Rijksuniversiteit Limburg (thans: Universiteit Maastricht), waaraan hij ook nu nog verbonden is. In de loop der jaren heeft hij vooral op het gebied van onderwijs (diverse aandachtsgebieden: epidemiologie, methoden en technieken, public health, gezondheidszorg ontwikkelingslanden, voeding), en in mindere mate op het gebied van onderzoek van zich doen spreken.

Martien volgde cursussen op het gebied van de epidemiologie bij o.a. Kleinbaum & Kupper en Miettinen. Zelf verzorgde hij ook de nodige cursussen, onder ander op basis van het in 1988 voor het eerst verschenen leerboek 'Epidemiologisch Onderzoek', waarvan hij samen met prof. dr. Lex Bouter (EMGO-instituut, VU Amsterdam) auteur is.

De in deze dissertatie beschreven studie naar de effectiviteit van *Ginkgo biloba* bij oudere mensen met geheugenklachten en beginnende dementie was de laatste jaren zijn belangrijkste activiteit in het kader van zijn onderzoekstaakstelling binnen de capaciteitsgroep Epidemiologie en het onderzoeksinstituut Nutrim.

